International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/31, 15/63, 15/67, 15/70, 15/74, 15/80, 15/82, 15/85, 15/11, 9/90, 1/19, 1/21, 5/10, C07K 14/37, A01K 67/027

A1

(11) International Publication Number: WO 00/32785

(43) International Publication Date: 8 June 2000 (08.06.00)

(21) International Application Number: PCT/IT99/00391

(22) International Filing Date: 29 November 1999 (29.11.99)

(30) Priority Data:
RM98A000730 30 November 1998 (30.11.98) IT

(71) Applicant *(for all designated States except US)*: UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA" [IT/IT]; Piazzale Aldo Moro, 5, I–00185 Roma (IT).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MACINO, Giuseppe [IT/IT]; Policlinico Umberto I, Università Degli Studi Di Roma "La Sapienza", Dipartimento Biotecnologie Cellulari ed Ematologia, Viale Regina Elena, 324, I–00161 Roma (IT). COGONI, Carlo [IT/IT]; Policlinico Umberto I, Università Degli Studi Di Roma "La Sapienza", Dipartimento Biotecnologie Cellulari ed Ematologia, Viale Regina Elena, 324, I–00161 Roma (IT).

(74) Agents: BANCHETTI, Marina et al.; Ing. Barzanò & Zanardo Roma S.p.A., Via Piemonte, 26, I–00187 Roma (IT).

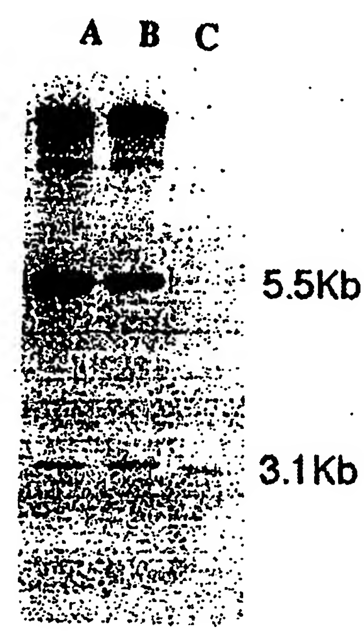
(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ISOLATION AND CHARACTERIZATION OF A *N. CRASSA* SILENCING GENE AND USES THEREOF

(57) Abstract

A nucleotide sequence encoding for a protein characterized in that it has a silencing activity and comprises a *recQ* helicase domain is disclosed; furthermore expression vectors suitable for the expression of said sequence in bacteria, plants, animals and fungi are disclosed; the invention refers also to organisms transformed by such vectors.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

ISOLATION AND CHARACTERIZATION OF A *N. CRASSA* SILENCING GENE AND USES THEREOF

The present invention relates to the isolation and characterization of a *Neurospora crassa* gene encoding for an essential activity in the co-suppression process and to uses and applications thereof in vegetal, animal and fungine fields.

The production of transgenic organisms is of large utility both in basic and applied biological research. The transgenic DNA is usually integrated in the genome and transferred as a Mendelian character. However, in various instances, the transgene introduction induces gene silencing phenomena (Flavell, R.B. 1994), i.e. the repression of the expression of the transgene itself and/or of one or more endogenous homologous genes.

The gene silencing can act at two levels: transcriptional (trans-inactivation) where transgenes contain sequences homologous to the silenced gene promoter (Vaucheret, 1993); and post-transcriptional (co-suppression) which requires homologies between coding regions (Flavell, 1994; Stam et al., 1997; Baulcombe, 1996).

Generally the silencing induced by a transgene requires an almost complete sequence homology (from 70% to 100%) between transgene and silenced gene sequences (Elkind, 1990).

In the *Neurospora crassa* filamentous fungus, during the vegetative phase, the presence of transgenes induces a post-transcriptional gene silencing phenomenon, named "quelling" (Cogoni et al., 1996).

By using the *al-1* gene (albino 1) (Schmidhauser et al., 1990) as silencing visual marker, many features of the phenomenon have been discovered (Cogoni et al., 1996). Particularly the *al-1* gene "quelling" in *Neurospora* is characterized in that: 1) the gene silencing is reversible further to the loss of transgene copies; 2) the reduction of mRNA basal level results from a post-transcriptional effect; 3) transgenes containing at least a region of 132 base pairs which is identical to the region encoding for the target gene are sufficient to induce the "quelling"; 4) the duplication of promoter sequences is ineffective to induce the silencing; 5) the "quelling" exhibits a dominant behavior in eterocarions containing both transgenic and untransformed nuclei, indicating the involvement of a molecule which is acts "in trans" among the nuclei; 6) the expression of an aberrant RNA transcribed by the transgenic locus is strictly correlated to silencing, suggesting that the "quelling" can be induced and/or mediated by a transgenic RNA molecule.

Therefore homologies between *Neurospora* silencing and plant co-suppression can be pointed out. The gene silencing in *Neurospora* is reversible, as result of transgenic copies instability during mitotic phase; in plants also the co-suppression reversion is associated with the reduction of transgene copy number, resulting from intra-chromosomal recombination during mitosis or meiosis (Mittelstein Scheid et al., 1994; Stam et al., 1998). Thus both in plants and in *Neurospora* the transgene presence is required to maintain the silencing. As in *Neurospora*, a decrease of the mRNA basal level of the silenced gene results from a post-transcriptional

mechanism (Dehio and Schell 1994; van Blokand et al., 1994; de Carvalho et al., 1995). Furthermore to induce the "quelling", transgenes must contain a portion of the silencing target gene coding sequence, being the promoter region ineffective. In plants coding regions with no promoter sequences can induce silencing (van Blokand et al., 1994) and, as in the "quelling", promoters or functionally active gene products are not required for the co-suppression.

One of the similarities between "quelling" and co-suppression in plants is that both mechanisms are mediated by diffusion factors. In *Neurospora* eterokaryotic strains, nuclei wherein the *albino-1* gene is silenced are able to induce the *al-1* gene silencing of the other not transformed nuclei, all sharing the same cytoplasmic environment (Cogoni et al., 1996). In plants the presence of a diffusion factor results from the fact that the co-suppression is effective in inhibiting the replication of Tobacco Etch Virus (TEV), a RNA virus with an exclusively cytoplasmic cycle. The occurrence of highly diffusible factors, which are effective to mediate the co-suppression, has been demonstrated using the grafting technique in tobacco (Palaqui et al., 1997), showing that silenced tobacco plants are able to transfer the silencing to non-silenced plants through grafting.

The fact that "quelling" and co-suppression share all these features suggests that mechanisms involved in post-transcriptional gene silencing in plants and in fungi can be evolved by an ancestral common mechanism.

Recently gene inactivation phenomena resulting from transgene introduction have been disclosed in animals. In *Drosophila melanogaster* the location of a transgene close

to heterochromatic centers results in a variegate expression (Wallrath and Elgin, 1995; Pirrotta, V., 1997). Similar expression profiles have been observed when the reference transgene is within tandem arrayed transposons, indicating that tandem repeats are effective to induce the chromatin condensation. (Dorer and Henikoff, 1994). Again in *Drosophila* Pal-Bhadra et al. (1997) have observed that the transgene introduction can lead to gene inactivation phenomena, similar to the co-suppression.

Gene silencing phenomena resulting from transegene sequence repeats have been disclosed recently in mammalians.

Garrick et al. (1998) produced mouse transgenic lines wherein 100 transgenic copies are present only in a locus and are directly tandem arrayed. The transgene expression has been disclosed to be inversely proportional to the number of occurring copies, indicating that silencing phenomena dependent on repeat copies are present also in mammalians.

Therefore the identification of *Neurospora* genes which are involved in the silencing is the first step to modulate the same process in plants, animals and fungi. The silencing modulation is of great relevance when transgenic organisms able to express the desired phenotype are produced.

The authors of the present invention have already isolated *Neurospora crassa* strains having mutations regarding essential functions for gene silencing mechanism (Cogoni and Macino, 1997); 15 independent isolated mutants define three complementation groups, thus identifying the *qde-1, qde-2 and qde-3* genes (*qde*

stands for "quelling"-deficient), whose products are essential to the silencing machinery. *qde* genes are essential to the *Neurospora* silencing, as suggested by the fact that silencing of three independent genes (*al-1*, *al-2* and *qa-2*) is impaired by *qde* mutations (Cogoni and Macino, 1997).

The authors of the invention have identified and cloned now one out of *Neurospora qde* genes, thus identifying one of required factors for silencing. By considering the similarity between "quelling" and co-suppression, genes orthologous to the isolated gene are involved in co-suppression and more generally in gene silencing in other organisms, like plants, fungi and animals.

The present invention can be applied with reference to two general scope: 1) silencing potentiation as a tool for inactivating more effectively and durably a desired gene, and 2) silencing suppression to obtain a better expression of the introduced transgenes.

As to the silencing potentiation, the over-expression of one or more genes controlling the phenomenon can lead to higher efficiency and/or stability thereof. Therefore the introduction of *qde-3* gene or of homologous genes thereof in microorganisms can constitute a tool to repress more effectively gene functions. Particularly this approach is specially useful in plants wherein the co-suppression is usually used for the "knock-out" of gene functions. In plants again the gene silencing potentiation can be used to obtain lines resistant to pathogen virus, by introducing transgenes encoding for viral sequences, in order to achieve the

expression inhibition of the virus itself (Flavell et al., 1994).

Analogous applications are suitable for animals, wherein some indications suggest that silencing can inhibit the suitable expression of introduced transgenes (Garrick et al., 1998).

On the contrary, there are instances wherein it is desirable not to have or to reduce the gene silencing, i.e. where a transgene is to be over-expressed. It is known that the co-suppression is strictly correlated both with the presence of an high copy number of the transgene, and with a transgene high expression. This correlation can hamper the production of transgenic organisms which express a transgene at high levels, because more high is the expression and/or the copy number, more probable is to evoke silencing responses. As above mentioned, analogous mechanisms of gene inactivation, dependent on a high copy number, have been disclosed in animals. In these circumstances plant or animal lines, totally or partially ineffective for silencing, constitute an ideal recipient wherein the desired gene can be over-expressed. The invention can be applied within this scope using different approaches:

A) Identification and production of mutant lines in genes homologous to *qde-3* gene, in plants, animals and fungi.

The knowledge of *Neurospora qde-3* gene, essential for silencing mechanism, can allow the isolation of mutant lines in other organisms, mutated in genes homologous to *qde-3*. For example by means of amplifications using degenerated primers, designed from the most conserved regions of *qde-3* gene, mutant lines in

homologous genes can be identified, by analysis of insertion mutant gene banks, already available for many plant species. Both in fungi and animals such mutants can be obtained, following the identification of the homologous gene, by means of "gene disruption" techniques using homologous recombination.

B) Reduction of *qde-3* gene expression

Other strategies for the production of silencing-deficient lines comprise the use of *Neurospora qde-3* gene or homologous genes thereof. *qde-3* or homologous genes can be introduced into suitable expression vectors to express them in an anti-sense orientation in order to inhibit the expression of resident endogenous genes. Alternatively portions of *qde-3* or of homologous genes can be over-expressed, in order to obtain a negative dominant effect and thus blocking the function of qde-3 endogenous genes.

The authors of the present invention have cloned and characterised the *Neurospora crassa qde-3 gene*. The sequence analysis showed that *qde-3* gene belongs to a highly conserved gene family, from *E. coli* to humans, named *recQ*. Genes belonging to this family encode for DNA helicase, as demonstrated by *in vitro* assays (Gray et al., 1997). The recQ helicase family is involved in recombinant processes. Mutations of these genes produce iper-recombinant phenotypes as, for example, the *S. cerevisiae* Sgs-1 gene involved both in meiotic and mitotic recombination.

The authors of the invention for the first time have demonstrated that a gene encoding for a recQ DNA-helicase is involved in gene silencing induced by transgenes. Therefore for the first time it is disclosed

that a gene belonging to the *recQ* family, other than acts during recombination, is also an essential component of the inactivation of repeat sequences.

Therefore it is an object of the invention a nucleotide sequence encoding for a protein characterized in having a silencing activity and comprising a recQ helicase domain, wherein the domain is at least 30% homologous with the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1. More preferably said homology is of at least of 60%. Most preferably the recQ helicase domain comprises the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1. According to a particular embodiment the nucleotide sequence encodes for a protein having the amino acid sequence of SEQ ID No. 1, or functional portions thereof. Even more preferably the nucleotide sequence of the invention is the sequence of SEQ ID No. 1 or its complementary sequence.

A further object of the invention is an expression vector comprising, under the control of a promoter that is expressed in bacteria, the nucleotide sequence of the invention. Those skilled in the art will appreciate that any plasmid suitable for a correct and effective expression of the protein of the invention in bacteria can be used and is within the scope of the invention.

A further object of the invention is an expression vector comprising, under the control of a promoter which is expressed in plants or in specific plant organs, the nucleotide sequence of the invention, both in a sense and anti-sense orientation. Those skilled in the art will appreciate that any plasmid suitable for a correct and effective expression of the protein of the invention in

plants or in specific plant organs can be used and is within the scope of the invention.

A further object of the invention is an expression vector comprising, under the control of a promoter which is expressed in fungi or in portions thereof, the nucleotide sequence of the invention, both in a sense and anti-sense orientation. Those skilled in the art will appreciate that any plasmid suitable for a correct and effective expression of the protein of the invention in fungi or in portions thereof can be used and is within the scope of the invention.

A further object of the invention is an expression vector comprising, under the control of a promoter that is expressed in animals, the nucleotide sequence of the invention both in a sense and anti-sense orientation. Those skilled in the art will appreciate that any plasmid suitable for a correct and effective expression of the protein of the invention in animals can be used and is within the scope of the invention.

A further object of the invention is a prokaryotic organism transformed by using the expression vector active in bacteria of the invention.

A further object of the invention is a plant or a specific plant organ transformed by using the expression vector active in plants of the invention.

A further object of the invention is a plant mutated at the nucleotide sequence of the invention and having a reduced or inhibited silencing activity.

A further object of the invention is a fungus transformed with the expression vector of the invention active in fungi.

A further object of the invention is a fungus mutated at the nucleotide sequence of the invention and having a reduced or inhibited silencing activity.

A further object of the invention is a non-human animal transformed with the expression vector of the invention active in animals.

A further object of the invention is a non-human animal mutated at the nucleotide sequence of the invention and having a reduced or inhibited silencing activity.

A further object of the invention refers to a protein characterized in having a silencing activity and in comprising a recQ helicase domain, wherein the domain is at least 30% homologous to the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1. Preferably the recQ helicase domain is at least 40% homologous with the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1. More preferably the recQ helicase domain is at least 60% homologous with the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1. Most preferably the recQ helicase domain comprises the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1. According to a particular embodiment the protein comprises the amino acid sequence of SEQ ID. No.1 or functional portions thereof.

It is within the scope of the invention the use of the nucleotide sequence of the invention to modulate gene silencing in plants, animals and fungi.

It is within the scope of the invention the use of the nucleotide sequence of the invention to potentiate the antiviral-response in a plant.

The present invention now will be disclosed by way of non limiting examples with reference to the following figures:

Figure 1: Southern blot analysis of genomic DNA extracted from (A): untransformed wild type strain, (B): 6xw recipient strain and (C): untransformed wild type strain, SmaI and HindIII digested, blotted and *al-1* gene probe hybridized. The 3.1-Kb band corresponds to the endogenous *al-1* gene, while the 5.5-Kb band corresponds to tandem arrayed *al-1* transgenes. The larger band represents undigested methylated DNA.

Figure 2: Linear map of the pMXY2 plasmid. Plasmid genes are shown as box. bmI: beta-tubulin allele which is responsible for benilate resistance; Amp: ampicillin resistance; qa-2 P: qa-2 gene promoter; TrpC T: trpC gene terminator. SphI and BglII are restriction sites used for the plasmid recovery from the 627 mutant chromosomal DNA.

Figure 3: Schematic representation of pQD6 and pQ35 plasmids. Restriction sites (BglII for pQD6 and SphI for pQ35) used for the recovery of the chromosomal DNA of the 627 strain are reported. Chromosomal sequences, flanking the integration site, are represented as segments. Restriction sites used to isolate DNA fragments used for probing the gene library are also represented.

Figure 4: Nucleotide sequence of the 6.9-Kb fragment containing the *qde-3* gene and flanking sequences. The amino acid sequence is shown above the nucleotide sequence. The bold sequences represent two introns of 98 and 68 nt. In these regions the underlined nucleotides identify consensus sequences of the donor site, the acceptor site and the internal sequence or lariat. It is also represented the pMXY2 plasmid

insertion site, in the 627 mutant, used for insertional mutagenesis. The portion encoding for the helicase domain is underlined.

Figure 5: Nucleotide sequence (SEQ ID No. 1) of the encoding portion reported in Figure 4 and deduced amino acid sequence. Amino acids from 897 to 1330, which define the recQ DNA-helicase domain, are underlined.

Figure 6: Multiple alignment, at the conserved domains, among *qde-3* and other proteins belonging to *recQ* family. arab *recQ*: A. *thaliana* isologous; *E. coli recQ; S. pombe hus-2; S. cerevisiae sgs-1;* human wrn: Werner syndrome; human blm: Bloom syndrome. Identical amino acids are shown in bold.

MATERIALS AND METHODS

E. coli strains

E. coli strain HB101 (F^-, hsdS20(rb^-, mb^-), supE44, recA13, ara14, proA2, rspL20(str^r), xyl-5) was used for cloning.

Neurospora crassa strains and growing conditions

Neurospora crassa following strains, supplied by Fungal Genetic Stock Center (FGSC, Dpt. Of Microbiology, University of Kansas Medical Ctr. Kansas City, KA) were used:

- Wild type (FGSC 987);
- qa-2/aro9 (FGSC 3957A), (FGSC 3958a).

The 6XW strain (Cogoni et al., 1996) was obtained upon transformation of the FGCS 3958a strain with pX16 (Cogoni et al., 1996). This plasmid contains the *qa-2* gene used as selective marker and the *al-1* coding sequence.

The mutated strains M7, M20 (*qde-1*); M10, M11 (*qde-2*); M17, M18 (*qde-3*) are described in Cogoni and Macino, 1997.

The *qde* mutants were obtained by UV mutagenesis. As recipient the transforming strain (6xw) silenced at the *albino-1* gene was used. *qde* mutants were selected for their ability to recover a wild type unsilenced phenotype and then classified in three different complementation groups. By analyzing the *al-2* gene quelling frequency all of *qde* used mutants are defective for the general silencing mechanism.

Complementation assays with not forced heterocaryons were carried out according to Davis and DeSerres, 1970.

Plasmids and libraries

The plasmid pMXY2, disclosed in Campbell et al., used for insertional mutagenesis was obtained from FGSC. The plasmid contains the *Bml* gene (allele responsible of the benilate drug resistance), that was used as selective marker after transformation. The genomic DNA containing the *qde-2* gene was isolated from a *N. Crassa* gene library in cosmids. (Cabibbo et al., 1991).

N. crassa transformation

Spheroplasts were prepared according to the Akins and Lambowitz (1985) protocol.

Southern Blot Analysis

Chromosomal DNA was prepared as disclosed by Irelan et al., 1993. 5 µg of genomic DNA were digested and blotted as reported in Maniatis et al.

DNA probes were: a) as to the *al-1* gene the probe is represented by a XbaI-ClaI restriction fragment of

pX16 (Cogoni et al., 1996); b) as to the *BmI* gene the probe is represented by the 2.6Kb SalI fragment of pMXY2.

Northern Blot Analysis

N. crassa total RNA was extracted according to the protocol described by Cogoni et al., 1996. The mycelium was grown for two days at 30°C, then powdered in liquid nitrogen before RNA extraction. For Northern analysis 10 µg of RNA were formaldehyde denatured, electrophoresed on a 1% agarose, 7% formaldehyde gel, and blotted over Hybond N (Amersham) membranes. Hybridization was carried out in 50% formamide in the presence of ^{32}P labeled DNA probe $1.5x10^{6}$ cpm/ml.

RESULTS

Isolation of silencing mutant by insertional mutagenesis

Neurospora strain (6XW) wherein the *albino-1* resident gene was steadily silenced was UV mutagenised, and *qde* ("quelling" deficient) mutants were isolated (Cogoni and Mancino 1997). The 6XW strain shows an albino phenotype due to the lack of carotenoid biosynthesis, as results by the silencing of the albino 1 gene expression (Schmidhauser et al., 1990). A mutation interfering with the silencing machinery is easily detectable by producing a wild type phenotype (bright orange) of the carotenoid biosynthesis. By means of complementation assays it was possible to establish that *qde* mutants belong to three complementation groups, indicating the presence of three genetic loci involved in the *Neurospora* silencing mechanism. In order to isolate the *qde* genes an insertional mutagenesis was carried out with the 6XW strain, previously used for UV mutagenesis. The insertional mutagenesis was carried out by transforming the 6XW strain with a plasmid, taking advantage of the

fact that, after the transformation, plasmids are randomly inserted in the *Neurospora crassa* genome. The mutagenesis was carried out transforming the 6XW silenced strain with pMXY2 (see Materials and Methods) which contains the benilate resistance as selective marker. Transformed strains able to grow in the presence of benilate containing medium and showing a wild type phenotype for the carotenoid biosynthesis were selected. Out of 50.000 isolated independent transformed strains, a benilate resistant strain (627) was isolated, which showed the bright orange phenotype expected for a *qde* gene mutation. In order to verify that the silencing release was effectively due to a *qde* gene mutation and not to the loss of *al-1*, the genomic DNA of the strain 627 was extracted and digested with SmaI and HindIII restriction enzymes. After blotting, DNA was hybridized with a probe corresponding to the coding sequence *of al-1.* The SmaI site is present only once in the *al-1* transgene containing plasmid and the digestion by using said enzyme produces a 5.5Kb fragment corresponding to tandem arrayed *al-1* transgenes, while a 3.1Kb fragment is expected from the resident *al-1* locus. Figure 1 shows that the number of *al-1* transgenic copies present in the 627 strain is comparable to that present in the silenced 6XW strain.

The 627 strain includes a mutated *qde3* gene

The 627 strain was assayed in a heterokaryon assay with a wild type strain and with M7, M20 (*qde-1*) M10, M11 (*qde-2*) mutants (Cogoni and Macino, 1997). As shown in Table 1 the *al-1* gene silencing is restored producing an albino phenotype in all of heterocaryons but M17 and M18.

This behavior is consistent with the presence of a *qde-3* gene recessive mutation in the 627 strain.

Table 1

Reciprocal heterokaryons among 627 mutant and previously characterized *qde* mutants.

	627	M7	M20	M10	M11	M17	M18
627	WT	AL	AL	AL	AL	WT	WT
M7		WT	WT	AL	AL	AL	AL
M20			WT	AL	AL	AL	AL
M10				WT	WT	AL	AL
M11					WT	AL	AL
M17						WT	WT
M18							WT

WT = heterokaryon with a wild type phenotype for carotenoid;

AL = heterokaryon with an albino phenotype wherein the al-1 gene silencing is restored.

<u>Recovery of sequences flanking the pMXY2 plasmid integration site</u>

In order to recover sequences flanking the integration site or sites the following methodology was carried out. The 627 strain genomic DNA was restricted with SphI and BglII enzymes. As shown in the map of Figure 2 the enzymes digest respectively upstream and downstream to the region containing both the ampicillin resistance gene and the origin of replication present in pMXY2. Subsequently the genomic DNA was ligated and the product used to transform *E. coli* cells. The screening was performed in an ampicillin-containing medium. pQD6 and pQ35 plasmids were recovered from BglII and SphI

restricted chromosomal DNA, respectively (see Figure 3). Two DNA fragments containing sequences flanking the integration site were isolated by using, respectively, BglII and SalI enzymes for pQD6, and SphI and HindIII enzymes for pQ35 (Figure 3).

Isolation of genomic clones, their subcloning and complementation of the *qde-3* mutant

The two fragments from pQD6 and pQ35 plasmids were used to probe a *Neurospora crassa* genomic library in cosmids. Cosmids 6E8 and 54D7, both containing about 30 Kb genomic DNA inserts, were isolated. Both the probes recognize the same cosmids, thus indicating that the two flanking sequences are contiguous. Cosmids 6E8 and 54D7 were used in transformation experiments with M17 and M18 mutants. Both of cosmids are able to restore the *al*-1 gene silencing in the two mutants, determining an albino phenotype. Furthermore the introduction of same cosmids into the M10 (*qde-2*) or the M20 mutant (*qde-1*) is not effective to restore the silencing.

The 6E8 cosmid was used to subclone a 9 Kb SphI-SphI fragment. This subclone was used for transformation experiments and resulted to be able to complement the *qde-3* phenotype, indicating that a *qde-3* functional gene is present in this plasmid.

Isolation and sequence of the qde-3 cDNA

The SphI-SphI region was sequenced, like the corresponding cDNA, by using RT-PCR. The latter sequence was used to deduce the qde-3 amino acid sequence and map the introns therein. The qde-3 gene encodes for a 1900 aa. putative protein (200 KDa). The genomic clone contains two introns of 98 nt. and 68 nt., respectively. Intron acceptor and donor sequences were identified and

correspond to described consensus sequences (Figure 4). Furthermore the pMXY2 plasmid insertion site within the gene in the 627 transforming strain is indicated. The insertion site was deduced by analysis of pQD6 and pQ35 plasmid sequences.

The cDNA sequence is shown in Figure 5 (SEQ ID No. 1), wherein the helicase domain containing 434 amino acids from 897 aa to 1330 aa is underlined.

The *qde-3* gene is belonging to recQ helicase DNA family

The 1900 aa sequence was used to search in database of amino acid sequences, by using the BLASTP algorithm. Significant homologies were identified with 6 genes belonging to the *reQ* family, belonging to the helicase group containing the DEAH consensus sequence. Figure 6 shows the homologous region sequence alignment of helicase domains, as defined in Figure 5, among *qde-3* and genes belonging to *recQ* helicase family. *qde-3* shows the highest homology with *hus-2* (55% amino acid identity) and the lowest homology with *Wrn* (40% identity).

Plant expression vector

The *qde-3* gene was inserted, in a sense orientation, into a vector containing a plant expression "cassette", including the 35S promoter and the PI-II "terminator" sequences. The vector also includes the *Streptomyces hygroscopicus bar* gene, which confers the phosphinotricine herbicide resistance to transformed plants. In an analogous vector, *qde-3* was inserted in an anti-sense orientation with respect to the 35S promoter.

The obtained vectors can be utilized to over-express the *qde-3* gene in plants, or to repress the gene expression of resident genes, which are homologous to *qde-3*, respectively.

Fungus expression vector

The *qde-3* gene was inserted in a vector containing a fungal specific expression "cassette", comprising the *A. nidulans trpC* gene promoter and terminator, both in a sense and an anti-sense orientation. In addition the vector contains the bacterial *hph* gene, which confers the hygromicine drug resistance. The sense plasmid can be used to over express the *qde-3* gene, whereas the anti-sense plasmid is used to repress the expression of *qde-3* homologous genes in various fungine species.

Mammalian expression vector

The *qde-3* gene was inserted in a vector containing a mammalian specific expression "cassette", including the cytomegalovirus (CMV) promoter and SV40 termination and polyadenylation sequences both in a sense and anti-sense orientation. The vector includes also the neomicine phototransferase gene, as marker for mammalian cell selection. The sense plasmid can be used to over express the *qde-3* gene, whereas the anti-sense plasmid can be used to repress the expression of *qde-3* homologous genes in various mammalian species.

Bibliography

- Akins, R.A. and Lambowitz A.M. (1985) Mol. Cell. Biol. 5:2272-2278
- Baulcombe, D.C. (1996) Plant Mol. Biol. 32, 79-88.
- Cogoni, C. et al. (1996) EMBO J. 15, 3153-3163
- Cogoni, C. and Macino, G. (1997) Proc. Natl. Acad. Sci. U.S.A. 94: 10223-10238.
- Cabibbo, A. et al. (1991) Fungal Genetic Newsl., 38: 68-70.
- Davis, R.H. and De Serres, F.J. (1970) methods Enzymol. 17: 79-143.

- de Carvalho Niebel, F. et al. (1995), Plant Cell : 347-358.
- Dehio, C., and Schell, J. (1994). Proc. Natl. Acad. Sci. U.S.A. 91: 5538-5542.
- Dorer, D.R. and Henikoff, S. (1994). Cell, 993-1002.
- Elkind, Y. Et al. (19990) Proc. Natl. Acad. Sci. U.S.A. 87: 9057-9061.
- Ellis, N.A. et al. (1995) Cell 83:655-666.
- Flavell, R.B. (1994) Proc. Natl. Acad. Sci. U.S.A. 91: 3490-3496.
- Garrick D., et al. (1998) Nature Genetics 18, 56-59.
- Gray, M.D. et al. (1997) Nature Genetics 17, 100-103.
- Irelan, J. et al. (1993) Fungal Genetics Newsl. 40: 24.
- Lindbo, J.A. et al. (1993) Plant Cell 5 : 1749-1759.
- Maniatis, S.T. et al. (1982) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Mittelstein Scheid, O. Et al. (1994) Mol. Gen. Genet. 244: 325-330.
- Orbach, M.J. et al. (1986) Mol. Cell. Biol., 6, 2452-2461.
- Pal-Bhadra, M., et al., (1997). Cell 90, 479-490.
- Palauqui, J.C. et al., (1997) EMBO J. 16: 4738-4745.
- Pirrotta, V. (1997). TIG 13, 314-318.
- Schmidhauser, T.J. et al., Mol. Cell. Biol. 10: 5064-5070
- Stam, M. et al. (1997) Annals of Botany 79:3-12
- Stewart, E. et al., (1997) EMBO J. 10:2682-2692
- van Blokland, R. et al. (1994), Plant, 6, 861-887.

- Vaucheret, H. (1993), C.R. Acad. Sci. Paris, Sciences de la vie/Life sciences 316, 1471-1483.
- Wallrath, L.L. and Elgin, S.C.R. (1995). Genes & Development 9, 1263-1277.
- Watt, P.M. et al. (1995). Cell, 81: 253-260.
- Watt, P.M. et al., (1996). Genetics 144: 935-945.
- Yu, C.E. et al. (1996). Science 272: 258-262.

Claims

1. Nucleotide sequence encoding for a protein characterized in having a silencing activity and in comprising a *recQ* helicase domain, wherein the domain is at least 30% homologous with the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1.

2. Nucleotide sequence encoding for a protein characterized in having a silencing activity and comprising a *recQ* helicase domain according to claim 1, wherein the domain is at least 40% homologous with the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1.

3. Nucleotide sequence encoding for a protein characterized in having a silencing activity and comprising a *recQ* helicase domain according to claim 2, wherein the domain is at least 60% homologous with the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1.

4. Nucleotide sequence encoding for a protein characterized in having a silencing activity and comprising a *recQ* helicase domain according to claim 3, wherein the recQ helicase domain is the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1.

5. Nucleotide sequence encoding for a protein characterized in having a silencing activity and comprising a *recQ* helicase domain according to claim 4, wherein said nucleotide sequence encodes for a protein having the amino acid sequence of SEQ ID No. 1, or functional portions thereof.

6. Nucleotide sequence encoding for a protein characterized in having a silencing activity and

comprising a *recQ* helicase domain according to claim 5, wherein said nucleotide sequence is the sequence of SEQ ID No. 1 or its complementary sequence.

7. Expression vector comprising, under the control of a promoter that is expressed in bacteria, the nucleotide sequence according to any one of claims 1-6.

8. Expression vector comprising, under the control of a promoter that is expressed in plants or in specific plant organs, the nucleotide sequence according to any one of claims 1-6, both in a sense and anti-sense orientation.

9. Expression vector comprising, under the control of a promoter that is expressed in fungi, the nucleotide sequence according to any one of claims 1-6 both in a sense and anti-sense orientation.

10. Expression vector comprising, under the control of a promoter that is expressed in animals, the nucleotide sequence according to any one of claims 1-6 both in a sense and anti-sense orientation.

11. Prokaryotic organism transformed by using the expression vector active in bacteria according to claim 7.

12. Plants or a specific plant organ transformed by using the expression vector active in plants according to claim 8.

13. Plant mutated at the nucleotide sequence according to any one of claims 1-6 having a reduced or inhibited silencing activity.

14. Fungus transformed by using the expression vector active in fungi according to claim 9.

15. Fungus mutated at the nucleotide sequence according to any one of claims 1-6 having a reduced or inhibited silencing activity.

16. Non-human animal transformed by using the expression vector active in animals according to claim 10.

17. Non-human animal mutated at the nucleotide sequence according to any one of claims 1-6 having a reduced or inhibited silencing activity.

18. Protein characterized in having a silencing activity and comprising a *recQ* helicase domain wherein the domain is at least 30% homologous to the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1.

19. Protein characterized in having a silencing activity and comprising a *recQ* helicase domain according to claim 18 wherein the domain is at least 40% homologous to the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1.

20. Protein characterized in having a silencing activity and comprising a *recQ* helicase domain according to claim 19 wherein the domain is at least 60% homologous to the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1.

21. Protein characterized in having a silencing activity and comprising a *recQ* helicase domain according to claim 20 wherein the domain is the amino acid sequence from aa. 897 to aa. 1330 of SEQ ID No.1.

22. Protein characterized in having a silencing activity and comprising a *recQ* helicase domain according to claim 21 comprising the amino acid sequence of SEQ ID. No.1 or functional portions thereof.

23. Use of the nucleotide sequence according to any one of claims 1-6 to modulate the gene silencing in plants, animals and fungi.

24. Use of the nucleotide sequence according to any one of claims 1-6 to potentiate the antiviral-response in a plant.

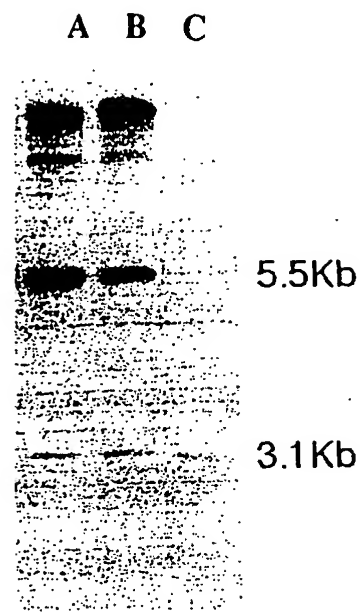
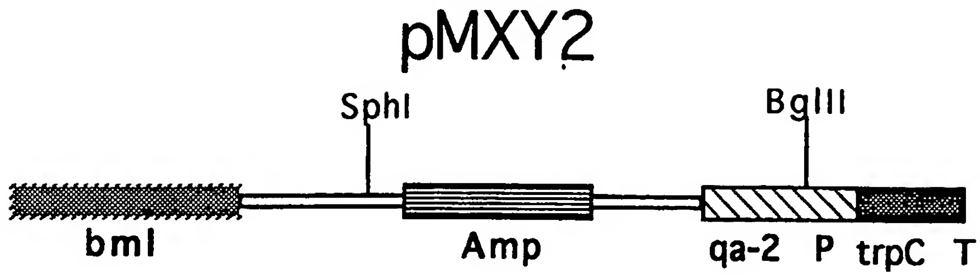


FIG. 1

FIG. 2



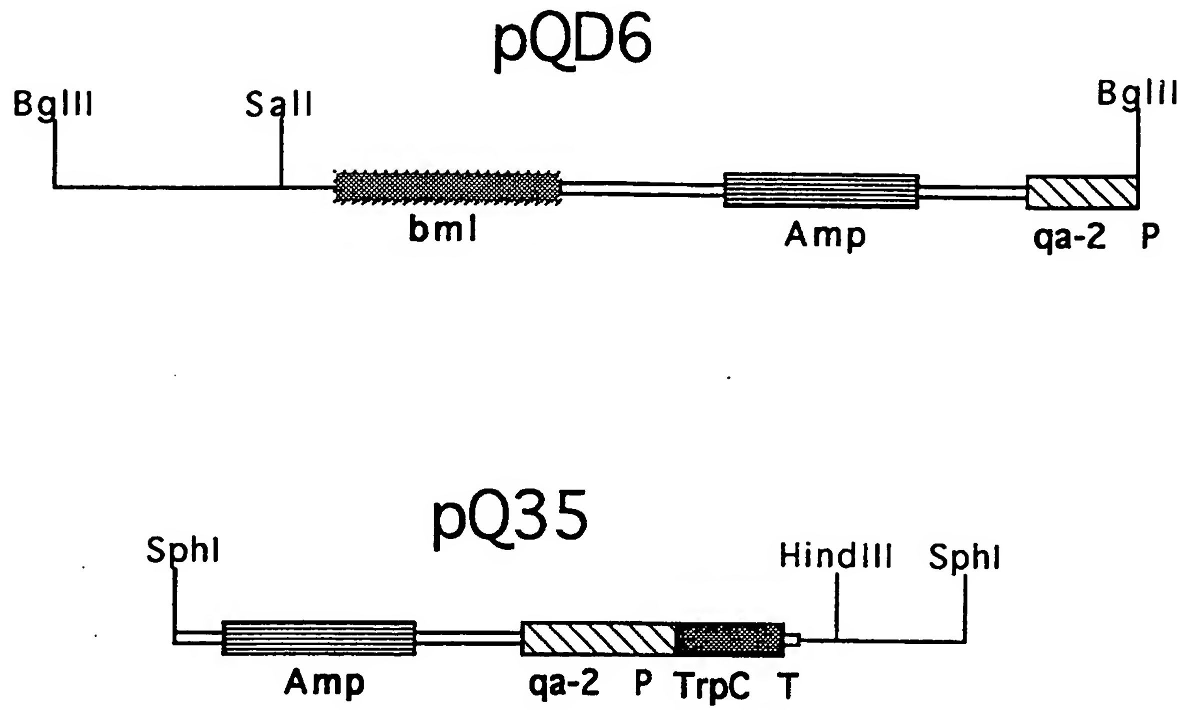

pQD6
BglII
SalI
BglII
bml
Amp
qa-2 P
pQ35
SphI
HindIII
SphI
Amp
qa-2 P
TrpC T

FIG. 3

```
CCA CCA CCA CCA CCA CCA CCA CAA CAA CCA ACT CAA CAA CGA ACT CAA CAA CCA
            9          18          27          36          45          54

ACC CCA AAC CCA ACC TCG ACC TCA ACC TCA ACC TCA ACC CTT GCG ACC TCG AGA TAC ACA
        63          72          81          90          99         108

AAC ACA TCC TCG ACA AAT GAC GCC CGA CCC GCT ACA CGC CAA CAG ATT GCC CCC GAG GTA
       123         132         141         150         159         168

                                                                 M   A
GGA GCA TCG ACG CAT CAG GAT TCA GTT GGA CTT GGA GAA GGA GGA GGA GGA GGA ATG GCG
       183         192         201         210         219         228

K   L   S   V   K   N   N   L   P   R   P   H   L   V   S   L   S   S   S   T
AAG CTC TCA GTC AAG AAC AAC CTG CCA CGG CCG CAC TTG GTC TCC TTG TCG TCG TCA ACG
       243         252         261         270         279         288

T   G   S   G   S   G   S   A   S   R   S   A   S   A   K   H   G   S   A   G
ACA GGC TCT GGG TCT GGT TCT GCG TCT AGG TCA GCT TCT GCT AAG CAC GGA AGT GCC GGT
       303         312         321         330         339         348

S   S   T   F   D   H   E   Q   H   Q   Q   H   Q   Q   Q   Q   Q   Q   K   R
TCC AGT ACC TTT GAT CAT GAA CAA CAT CAA CAA CAT CAA CAA CAA CAA CAA CAA AAG CGC
       363         372         381         390         399         408

Q   R   S   Q   S   E   A   R   Q   Q   Q   Q   Q   Q   Q   Q   Q   Q   Q   Q
CAG CGG TCG CAA TCA GAA GCA CGA CAA CAG CAG CAG CAA CAG CAA CAG CAA CAG CAA CAG
       423         432         441         450         459         468

Q   Q   Q   Q   Q   Q   A   Q   H   H   A   H   S   T   Y   A   Q   R   P   Q
CAA CAA CAA CAA CAA CAA GCA CAG CAC CAT GCA CAT TCT ACA TAT GCA CAA AGA CCC CAA
       483         492         501         510         519         528

P   T   P   Q   Q   R   P   P   Q   N   L   L   T   P   A   S   T   T   G   A
CCC ACC CCC CAA CAA CGA CCA CCC CAA AAC CTA CTG ACA CCT GCT TCA ACC ACT GGT GCC
       543         552         561         570         579         588

S   V   G   P   L   Q   R   A   Y   S   V   S   L   A   A   R   Q   S   P   S
AGC GTC GGC CCG CTC CAA CGC GCA TAC TCG GTT TCA TTA GCT GCG AGA CAG TCC CCC TCG
       603         612         621         630         639         648

T   N   L   V   R   P   K   T   D   S   P   A   P   H   T   L   H   L   K   N
ACA AAC TTG GTC CGT CCA AAG ACC GAC TCG CCA GCT CCC CAC ACT TTA CAC CTC AAG AAC
       663         672         681         690         699         708

K   K   N   L   R   H   P   A   P   T   P   D   S   P   I   V   D   D   D   I
AAG AAG AAC CTC CGT CAC CCC GCC CCC ACG CCC GAC AGT CCG ATC GTA GAC GAC GAT ATT
       723         732         741         750         759         768

F   S   D   A   V   D   L   T   E   E   L   D   H   D   H   D   L   N   G   K
TTC TCC GAC GCC GTC GAT CTT ACC GAA GAA CTC GAT CAT GAC CAT GAT CTC AAC GGC AAA
       783         792         801         810         819         828

D   K   D   N   T   D   N   D   N   T   V   A   S   S   S   L   I   G   F   G
GAC AAA GAC AAC ACC GAC AAC GAC AAC ACA GTC GCT TCC AGT TCG CTA ATA GGG TTC GGC
       843         852         861         870         879         888

D   D   K   L   L   W   R   E   D   F   A   E   R   A   E   P   E   H   E   R
GAT GAC AAG TTA CTG TGG CGA GAG GAC TTT GCT GAG CGT GCA GAG CCC GAA CAT GAA AGA
       903         912         921         930         939         948

G   G   S   R   P   R   Q   V   K   K   R   K   I   S   N   D   Y   I   M   K
GGT GGG AGC AGG CCT CGC CAG GTC AAG AAA CGG AAG ATA TCG AAT GAC TAC ATT ATG AAG
       963         972         981         990         999        1008

D   E   D   V   S   L   F   D   D   D   G   E   E   D   E   F   M   D   I   N
GAT GAG GAT GTC TCG CTT TTT GAT GAT GAT GGC GAG GAG GAC GAG TTT ATG GAT ATC AAT
      1023        1032        1041        1050        1059        1068
```

FIG. 4/1

```
E   L   V   Q   G   D   R   E   S   T   P   K   P   K   A   T   S   R   S   V
GAG CTA GTT CAG GGG GAT CGG GAA AGT ACT CCG AAG CCA AAG GCT ACA TCG AGG TCT GTC
       1083        1092        1101        1110        1119        1128

S   T   R   L   P   P   T   V   S   L   Q   R   G   R   S   P   K   R   K   E
TCG ACG AGG CTG CCG CCT ACA GTA TCG CTG CAA CGG GGT CGG TCT CCT AAG AGG AAG GAG
       1143        1152        1161        1170        1179        1188

A   S   V   E   K   R   T   T   E   N   Q   Q   Q   A   D   R   E   D   E   P
GCT TCA GTT GAA AAG CGC ACA ACG GAA AAC CAG CAA CAG GCT GAC AGA GAA GAC GAA CCG
       1203        1212        1221        1230        1239        1248

S   F   M   S   S   P   D   V   D   N   S   R   K   R   K   S   S   G   S   P
TCG TTT ATG TCA AGT CCA GAT GTC GAC AAC TCC CGC AAG CGA AAG TCT TCT GGA TCG CCC
       1263        1272        1281        1290        1299        1308

T   G   L   T   T   P   R   P   Q   Q   K   Q   T   E   E   V   P   G   T   T
ACA GGT TTA ACG ACG CCA AGA CCC CAG CAG AAG CAA ACG GAA GAG GTC CCA GGT ACG ACC
       1323        1332        1341        1350        1359        1368

T   A   K   K   P   R   R   S   E   V   M   D   S   E   D   E   A   F   T   P
ACC GCC AAG AAG CCA CGG CGC AGT GAA GTG ATG GAC TCG GAG GAC GAG GCA TTC ACT CCT
       1383        1392        1401        1410        1419        1428

L   S   A   G   S   L   P   G   S   A   E   F   F   R   S   G   G   T   T   T
CTT TCT GCT GGG TCG CTG CCT GGG AGT GCG GAG TTC TTC AGA AGC GGT GGG ACC ACC ACA
       1443        1452        1461        1470        1479        1488

R   E   L   G   L   D   E   D   T   V   M   D   T   P   S   R   P   P   V   E
CGG GAA TTG GGT TTG GAC GAA GAC ACG GTT ATG GAC ACG CCT AGT AGG CCA CCG GTC GAG
       1503        1512        1521        1530        1539        1548

S   T   L   P   T   L   E   S   V   E   S   R   P   P   P   L   P   P   M   D
TCC ACT TTG CCA ACT CTC GAG TCT GTG GAA AGT CGA CCA CCC CCC CTG CCG CCC ATG GAT
       1563        1572        1581        1590        1599        1608

L   P   S   Q   R   K   P   L   E   P   L   N   T   P   R   N   Q   L   L   E
CTA CCA TCA CAG CGA AAA CCG CTA GAG CCG TTG AAC ACT CCG CGC AAC CAG CTG CTT GAG
       1623        1632        1641        1650        1659        1668

S   V   E   R   P   T   Q   Q   P   S   V   G   P   S   F   A   Q   S   S   T
TCG GTC GAA AGG CCA ACA CAG CAG CCG TCG GTG GGG CCG AGT TTT GCA CAG AGT AGC ACA
       1683        1692        1701        1710        1719        1728

L   A   E   S   S   L   P   P   S   M   P   P   P   S   E   D   P   L   N   T
CTC GCC GAA AGC TCC CTG CCG CCG TCA ATG CCG CCG CCA AGT GAA GAC CCC CTC AAC ACC
       1743        1752        1761        1770        1779        1788

R   E   N   S   N   L   E   E   F   D   Y   K   L   Y   K   P   L   L   D   L
AGG GAG AAC AGC AAC CTT GAG GAG TTC GAC TAC AAG CTT TAC AAA CCC CTG CTA GAT CTT
       1803        1812        1821        1830        1839        1848

F   V   N   A   P   A   I   L   E   R   E   L   S   A   V   N   D   E   L   Q
TTC GTC AAC GCA CCC GCA ATC TTG GAA AGA GAA CTG AGC GCC GTT AAT GAC GAG CTT CAG
       1863        1872        1881        1890        1899        1908

E   N   M   I   K   L   R   D   C   L   R   L   P   R   E   E   R   D   R   A
GAG AAC ATG ATC AAG CTG CGG GAC TGT CTG CGC CTG CCC AGG GAA GAA AGA GAC AGG GCA
       1923        1932        1941        1950        1959        1968

R   E   E   V   K   K   E   K   E   M   L   K   R   R   D   I   A   L   R   A
CGC GAA GAG GTG AAG AAG GAA AAG GAA ATG CTC AAG CGA CGG GAC ATT GCG CTC AGA GCC
       1983        1992        2001        2010        2019        2028

L   Q   D   E   H   K   L   Y   V   K   K   R   K   E   H   N   L   I   N   E
CTC CAG GAC GAA CAC AAG TTG TAC GTC AAG AAA CGC AAA GAG CAT AAT TTG ATC AAC GAG
       2043        2052        2061        2070        2079        2088

E   I   V   R   A   Y   A   E   E   D   D   E   Y   E   D   Q   L   M   A   Q
GAA ATC GTT CGC GCT TAT GCT GAA GAA GAC GAT GAG TAC GAG GAT CAG TTA ATG GCG CAG
       2103        2112        2121        2130        2139        2148
```

FIG. 4/2

L D K L D D E V E A I V K S L T R L I V
CTG GAC AAG TTG GAT GAT GAG GTT GAG GCT ATC GTA AAG AGT CTG ACA AGG CTT ATT GTG
2163 2172 2181 2190 2199 2208

A A G I T E K S F D L K K E E E E E E E
GCG GCG GGG ATC ACG GAG AAG AGC TTT GAC CTA AAG AAG GAG GAG GAA GAG GAG GAG GAG
2223 2232 2241 2250 2259 2268

K P I I I A T P T P S T R T E A P V L P
AAG CCG ATC ATC ATA GCG ACT CCG ACA CCT TCG ACG AGG ACC GAG GCC CCG GTT CTG CCG
2283 2292 2301 2310 2319 2328

T T E Y H N S Q Q V I L Q T Q H P A A Q
ACG ACC GAG TAT CAT AAT TCC CAG CAG GTC ATA TTG CAG ACT CAA CAT CCT GCT GCG CAG
2343 2352 2361 2370 2379 2388

Q V S H R V P P P P T P S F Q T A R Q T
CAG GTT TCT CAC CGG GTG CCA CCA CCT CCG ACA CCG AGT TTT CAA ACG GCG CGC CAG ACT
2403 2412 2421 2430 2439 2448

P V S Y Q S R P T N N S F P D I S A E E
CCG GTG TCA TAT CAG AGC AGA CCG ACC AAC AAC TCC TTT CCT GAT ATC TCG GCG GAA GAA
2463 2472 2481 2490 2499 2508

A M M F D K E D P F M E Q Q H A P A S A
GCC ATG ATG TTC GAT AAA GAA GAC CCC TTC ATG GAA CAA CAG CAC GCC CCG GCC TCT GCT
2523 2532 2541 2550 2559 2568

P F Q A T L P Q R N S P F K T A P F K P
CCC TTC CAG GCC ACC CTT CCC CAG CGC AAC AGC CCT TTC AAA ACC GCC CCG TTC AAG CCA
2583 2592 2601 2610 2619 2628

V H G H D Y F D D E D D D A D L L A A V
GTC CAC GGC CAC GAT TAC TTT GAC GAT GAA GAC GAC GAT GCC GAC CTC CTG GCA GCA GTA
2643 2652 2661 2670 2679 2688

D S A E T Y T S T A A T T T T N N N N H
GAC AGC GCC GAG ACG TAT ACT TCT ACG GCC GCC ACC ACC ACC ACC AAC AAC AAC AAT CAC
2703 2712 2721 2730 2739 2748

L R S Q S V M S T S T A T T I K P R K R
TTA CGA TCA CAA TCG GTG ATG TCA ACA TCC ACG GCG ACC ACG ATC AAA CCG AGG AAA CGC
2763 2772 2781 2790 2799 2808

N E N A N A K K P K S V H A K L S M P P
AAC GAA AAT GCC AAT GCC AAG AAG CCC AAG TCC GTA CAT GCA AAG TTA TCG ATG CCG CCC
2823 2832 2841 2850 2859 2868

E K M K Y A W S N D V R K A L K D R F R
GAA AAG ATG AAG TAT GCG TGG TCG AAT GAT GTG AGG AAG GCT CTC AAG GAT AGG TTT CGG
2883 2892 2901 2910 2919 2928

M S G F R Q N Q L E A I N A T L G G K
ATG TCG GGG TTC AGA CAG AAT CAG TTG GAG GCT ATT AAT GCT ACT TTG GGT GGT AAG GTG
2943 2952 2961 2970 2979 2988

AGT TCT CTG TCC TTT ACC TAT CTG GGA GAG ACC AAG AAG GAG AGA GAG AGA GAG AGG AGG
3003 3012 3021 3030 3039 3048

D A F V L M P T G
GGA AGA CGA AAA TGG ACT TTG CTG ACT CTA GAAAG GAT GCC TTT GTG TTG ATG CCG ACT GGT
3065 3074 3083 3092 3101 3110

G G K S L C Y Q L P A V V R S G K T R G
GGT GGA AAG TCT CTG TGC TAT CAG TTG CCG GCT GTA GTC AGG AGC GGC AAG ACG CGT GGT
3125 3134 3143 3152 3161 3170

I T V V I S P L L S L M L D Q V N H L A
ATC ACA GTC GTC ATC TCC CCT CTG CTA AGT CTG ATG CTG GAT CAA GTC AAC CAT TTG GCA
3185 3194 3203 3212 3221 3230

FIG. 4/3

```
N   L   M   I   Q   A   Y   A   F   N   G   D   M   N   S   E   M   R   R   M
AAC CTG ATG ATC CAA GCT TAC GCT TTC AAC GGA GAC ATG AAC TCA GAA ATG CGC CGA ATG
        3245        3254        3263        3272        3281        3290

V   F   Q   K   L   D   A   E   H   P   E   H   E   L   Q   L   L   Y   V   T
GTG TTT CAG AAG CTT GAT GCT GAG CAT CCT GAG CAT GAG CTC CAA CTG CTC TAT GTC ACC
        3305        3314        3323        3332        3341        3350

P   E   M   V   S   K   N   Q   T   F   V   N   K   M   M   D   L   Y   R   R
CCG GAG ATG GTG AGC AAG AAC CAG ACA TTC GTC AAC AAG ATG ATG GAC CTC TAC CGG AGG
        3365        3374        3383        3392        3401        3410

K   K   L   A   R   I   V   I   D   E   A   H   C   V   S   Q   W   G   H   D
AAA AAG CTG GCT AGA ATT GTT ATC GAC GAG GCT CAC TGC GTC AGT CAA TGG GGC CAT GAC
        3425        3434        3443        3452        3461        3470

F   R   P   D   Y   K   A   I   G   E   F   R   K   R   F   P   G   V   P   V
TTC CGA CCC GAT TAC AAA GCT ATA GGA GAG TTT CGT AAG AGG TTT CCC GGA GTT CCG GTC
        3485        3494        3503        3512        3521        3530

M   A   L   T   A   T   A   T   Q   N   V   I   L   D   V   K   H   N   L   A
ATG GCT TTG ACA GCG ACG GCA ACA CAG AAC GTC ATC TTG GAT GTC AAG CAT AAC CTG GCA
        3545        3554        3563        3572        3581        3590

M   E   D   C   Q   T   F   S   Q   S   F   N   R   P   N   L   Y   Y   E   V
ATG GAG GAC TGC CAG ACT TTC TCC CAG AGC TTT AAT CGG CCG AAC CTC TAC TAT GAG GTC
        3605        3614        3623        3632        3641        3650

R   M   K   E   Q   N   L   I   A   R   I   A   E   L   I   K   E   K   Y   D
AGG ATG AAG GAG CAG AAT CTG ATT GCC CGC ATC GCA GAG TTG ATC AAG GAG AAG TAT GAC
        3665        3674        3683        3692        3701        3710

G   Q   T   G   I   I   Y   T   L   S   R   K   S   A   E   N   I   A   K   N
GGC CAG ACG GGT ATC ATC TAC ACA TTA TCA AGA AAG AGT GCC GAG AAC ATC GCC AAA AAT
        3725        3734        3743        3752        3761        3770

L   Q   E   K   H   R   I   K   A   K   H   Y   H   A   S   I   T   T   D   E
CTC CAG GAA AAA CAC CGC ATC AAA GCA AAG CAC TAC CAT GCC TCC ATC ACC ACC GAC GAA
        3785        3794        3803        3812        3821        3830

K   I   S   V   Q   H   E   W   Q   T   G   R   V   K   V   V   V   A   T   I
AAG ATC AGC GTT CAA CAT GAA TGG CAA ACC GGC CGA GTC AAA GTC GTG GTA GCC ACC ATT
        3845        3854        3863        3872        3881        3890

A   F   G   M   G   I   D   K   P   D   V   R   F   V   I   H   Q   H   I   P
GCC TTC GGC ATG GGC ATC GAC AAG CCT GAC GTC CGC TTT GTT ATC CAC CAG CAC ATC CCC
        3905        3914        3923        3932        3941        3950

K   S   L   E   G   Y   Y   Q   E   T   G   R   A   G   R   D   G   K   P   S
AAG TCG CTC GAA GGT TAC TAT CAA GAA ACC GGC CGC GCC GGA CGT GAC GGC AAG CCA TCG
        3965        3974        3983        3992        4001        4010

D   C   Y   L   Y   F   A   Y   G   D   I   Q   S   L   R   R   M   I   A   D
GAC TGC TAC TTG TAC TTT GCC TAT GGC GAC ATT CAA TCC CTA CGT CGT ATG ATC GCC GAC
        4025        4034        4043        4052        4061        4070

G   E   G   D   Y   A   Q   K   E   R   Q   L   Q   M   L   N   R   V   V   S
GGC GAA GGT GAC TAC GCG CAA AAG GAG CGT CAG CTA CAA ATG CTC AAC CGT GTG GTC AGC
        4085        4094        4103        4112        4121        4130

Y   C   E   S   Q   H   T   C   R   R   E   E   V   L   R   Y   F   G   E   E
TAT TGC GAG TCG CAG CAC ACG TGC CGG CGC GAA GAA GTG CTC CGC TAC TTT GGC GAG GAG
        4145        4154        4163        4172        4181        4190

F   D   Y   R   K   C   R   D   G   C   D   N   C   R   N   G   R   I   S   K
TTT GAC TAC CGG AAG TGT AGA GAC GGA TGC GAT AAC TGC CGG AAC GGA CGC ATC TCG AAG
        4205        4214        4223        4232        4241        4250

S   T   E   M   R   D   F   T   E   I   A   F   A   A   I   E   V   V   K   S
TCG ACG GAG ATG AGG GAT TTT ACG GAA ATC GCC TTC GCC GCG ATC GAG GTG GTG AAG AGC
        4265        4274        4283        4292        4301        4310
```

FIG 4/4

Q Q P I T L G K L C D I L M G K R K N E
CAG CAG CCC ATC ACG CTG GGC AAG CTG TGC GAC ATC CTG ATG GGC AAG AGA AAG AAC GAG
4325 4334 4343 4352 4361 4370

H G G V C H F G I A K G S T Q R E L Q R
CAC GGT GGC GTG TGT CAC TTT GGT ATC GCC AAG GGG AGC ACG CAG AGG GAG CTG CAG AGG
4385 4394 4403 4412 4421 4430

I V L Q L N F H K A L G E D N I M N G A
ATC GTG CTG CAG CTG AAT TTC CAC AAG GCG CTG GGC GAG GAC AAT ATC ATG AAT GGG GCG
4445 4454 4463 4472 4481 4490

G M P I T Y Y I
GGG ATG CCT ATT ACC TAC TAT ATT **GTG AGT GCT GTC CCG GTT GGT CTT GCA TAT CTG GCT**
4505 4514 4523 4532 4541 4550

/INSERZIONE
T G P E A G A Y L
TTG TTG CTT TGC TAA CAC AGC AGC TCG TAC AG ACC GGC CCT GAA GCT GGT GCT TAC CTC
4564 4573 4582 4591 4600 4609

Y N G K R L M L P V P S N K S V E P P S
TAC AAT GGC AAG CGG TTG ATG CTG CCA GTT CCC TCA AAC AAG TCC GTC GAA CCC CCG TCT
4624 4633 4642 4651 4660 4669

R S K Q R S R R V D E D M D E Q E L S T
CGG TCT AAG CAG CGG AGC CGT CGA GTC GAT GAG GAT ATG GAT GAG CAA GAA CTT TCC ACC
4684 4693 4702 4711 4720 4729

L Q R P P T S T N V S S P V R A T K K R
CTG CAA CGA CCG CCA ACA TCA ACA AAT GTC TCT TCA CCC GTT CGA GCC ACC AAG AAA CGA
4744 4753 4762 4771 4780 4789

S S K K A L P T L I A D Y E E P S S D G
AGT TCC AAA AAG GCT TTA CCG ACC CTC ATC GCC GAC TAC GAA GAG CCC AGC TCC GAC GGT
4804 4813 4822 4831 4840 4849

P H G P L H A N G Y E R D N F V V S D N
CCT CAC GGT CCT CTC CAC GCC AAC GGC TAT GAG CGC GAT AAC TTT GTC GTA TCC GAT AAT
4864 4873 4882 4891 4900 4909

V E P E E E E D A F E P V R P S R R G P
GTT GAA CCC GAA GAG GAA GAA GAT GCC TTC GAA CCT GTC CGC CCC TCG CGG CGC GGC CCA
4924 4933 4942 4951 4960 4969

S S R A T R P Q H R Q T T L Y D T L S H
TCT TCT CGC GCT ACC CGC CCT CAA CAC CGC CAG ACC ACC CTT TAT GAC ACC CTC TCC CAC
4984 4993 5002 5011 5020 5029

T Q Q S Q T V S Q H L A T L G P P I D A
ACC CAA CAA TCC CAA ACC GTC TCC CAA CAC CTC GCC ACT TTG GGT CCG CCC ATC GAC GCC
5044 5053 5062 5071 5080 5089

R T M H N P R Y A Q L D E V H Q D I V D
CGC ACC ATG CAT AAC CCC CGC TAC GCC CAG CTT GAC GAG GTC CAC CAG GAT ATT GTC GAT
5104 5113 5122 5131 5140 5149

A F V E E V K V F E E D F R N R N H M R
GCC TTT GTT GAA GAA GTC AAG GTC TTC GAG GAG GAC TTT CGC AAC AGG AAC CAC ATG CGC
5164 5173 5182 5191 5200 5209

K P I F T E T Q Y R E M A I R W T R S L
AAA CCC ATC TTT ACC GAG ACG CAG TAC CGT GAG ATG GCA ATC CGG TGG ACG CGG TCG TTA
5224 5233 5242 5251 5260 5269

D A M R A I P D I N Q D K V D R Y G A K
GAC GCG ATG CGC GCG ATC CCG GAT ATC AAC CAG GAT AAA GTA GAT CGG TAT GGT GCC AAA
5284 5293 5302 5311 5320 5329

F I P L V E R F W G N Y Q E M M G G G Y
TTC ATC CCA CTT GTG GAG CGG TTC TGG GGG AAT TAT CAG GAG ATG ATG GGG GGA GGG TAT
5344 5353 5362 5371 5380 5389

FIG 4/5

D N P A V A G D E D D D E G P R R T G N
GAT AAT CCT GCT GTG GCT GGC GAT GAG GAT GAT GAT GAG GGC CCC AGG AGG ACA GGA AAT
5404 5413 5422 5431 5440 5449

G K G G N K K G G G G G G G N E V V D L
GGA AAA GGG GGG AAT AAG AAG GGA GGA GGA GGA GGA GGA GGA AAT GAA GTA GTG GAT TTG
5464 5473 5482 5491 5500 5509

I S S D E D E P P A R A P S R N A G R G
ATT AGT AGT GAT GAG GAT GAA CCT CCG GCT CGT GCA CCA TCG CGG AAT GCG GGG CGA GGA
5524 5533 5542 5551 5560 5569

K A Q S T R G G Q I Q D K G R A V N R R
AAG GCA CAG TCG ACA CGT GGG GGA CAA ATC CAA GAT AAA GGC CGA GCA GTC AAC CGC CGC
5584 5593 5602 5611 5620 5629

G E P I A E E D E E D Y G L S D P D I D
GGA GAA CCC ATC GCC GAA GAA GAC GAA GAA GAC TAC GGG CTA AGC GAC CCC GAT ATC GAC
5644 5653 5662 5671 5680 5689

A I D P D A I T A S D N S D E E D D D D
GCC ATC GAT CCA GAC GCC ATC ACC GCC TCC GAC AAC TCC GAC GAA GAA GAT GAT GAT GAT
5704 5713 5722 5731 5740 5749

D D E D L E S S R Y F S G S T G P P V S
GAT GAC GAA GAC CTC GAA TCC TCC CGC TAC TTC TCC GGC TCA ACA GGC CCG CCC GTC TCC
5764 5773 5782 5791 5800 5809

K A V Q D A R L R E Q L S M Y A S G G S
AAA GCC GTG CAG GAT GCT CGA CTC CGT GAA CAA CTT TCC ATG TAC GCC TCC GGC GGC AGC
5824 5833 5842 5851 5860 5869

S S K G S Y G S G R A S G G S S S R A S
TCT TCG AAA GGT AGC TAC GGC TCA GGG CGC GCA TCA GGA GGA TCT TCG TCG AGA GCG TCG
5884 5893 5902 5911 5920 5929

G S G W R G G G A G G K K Y Y R K K R A
GGA TCA GGA TGG AGA GGT GGA GGA GCA GGT GGG AAG AAA TAC TAC AGG AAG AAG AGG GCT
5944 5953 5962 5971 5980 5989

G S S A A G G G G A G G G G V T K R K A
GGT TCT TCG GCT GCT GGT GGT GGT GGT GCA GGA GGA GGG GGA GTT ACA AAA CGG AAG GCG
6004 6013 6022 6031 6040 6049

S G S G A K T A R K R G A S T A P K T T
AGT GGG AGT GGC GCG AAG ACG GCG AGG AAG AGG GGT GCA TCT ACT GCG CCG AAG ACA ACG
6064 6073 6082 6091 6100 6109

T R G G G S G A G S R G G G A G G A G G
ACG AGA GGG GGA GGA TCT GGA GCT GGG TCT AGA GGA GGC GGT GCT GGT GGT GCT GGT GGT
6124 6133 6142 6151 6160 6169

A G A G A G A G G G K R G G G G G G G M
GCT GGT GCT GGT GCT GGT GCT GGA GGA GGG AAA AGG GGT GGT GGA GGT GGA GGA GGA ATG
6184 6193 6202 6211 6220 6229

G G I S V M P H .
GGA GGG ATA AGT GTT ATG CCT CAT TAG CTA TTT TAT AGC ATA TCG CAT TTA TAC AGT GTC
6244 6253 6262 6271 6280 6289

TTA TGG AAG GGA GGA GGA GAA GAA GAA GGA TAA GCT GGC ATA AGC TTG AAC CGG CCA GGC
6304 6313 6322 6331 6340 6349

CAA AAT GGC CAG AGA GCT CAC CGG GCA ATC GAG CTT GAA ATG AGC TTG ACA TAT TAG GTA
6364 6373 6382 6391 6400 6409

TTC CCG AGA ATA TAG CGG GAT TAC AAG GCA CTT ACT TTA CCA AGT CGA AAG GGA CGA GCC
6424 6433 6442 6451 6460 6469

AAA TCT ATG GTA CTC GCC AGT TGC GCA ACG TTG AGT TTT ATC ATT CGT GGA GTT TTC ATC
6484 6493 6502 6511 6520 6529

FIG. 4/6

GTG GAG TTT TTA TTA TCA ACT ATT CGT TGT ATA GTT TTC GTT GTA GAT GTT AGT TCC GGA
6544 6553 6562 6571 6580 6589

CGA TCA AAA GGG GAA GTG TGG AAC AGA GAA GTC GAA AGG ACA AGC CAA AAT GAC ATG GCA
6604 6613 6622 6631 6640 6649

GTG TCC AGT CAG ATA CCC TCC AGA CAA AAC CAG ACA CCA ATA ACA AAC CCT TCA ACC ATA
6664 6673 6682 6691 6700 6709

ACA CCA GCA AAG CCA ATC CTT AGG TAC CTA CCT AGG GTA GGG TAG GTC CAG GAA TGT CTT
6724 6733 6742 6751 6760 6769

CCC CAA AGG TAC CTC TAC TTA TTC ATG TTA CGC TCC ATC AGT CCC ATC GCT TAG CAT CGC
6784 6793 6802 6811 6820 6829

TGC CCG GTT ACC TAT CTC TAC CTC TAC CTC TAC CTC TAC CTC TAC CTC TAC CTC TAC CTC
6844 6853 6862 6871 6880 6889

TAT CTC TAC CTC TAC CTC TAC
6904 6913

FIG. 4/7

SEQ ID No.1

		M	A	K	L	S	V	K	N	N	L	P	R	P	H	L	V	S	L
		ATG	GCG	AAG	CTC	TCA	GTC	AAG	AAC	AAC	CTG	CCA	CGG	CCG	CAC	TTG	GTC	TCC	TTG
				9			18			27			36			45			54
S	S	S	T	T	G	S	G	S	G	S	A	S	R	S	A	S	A	K	H
TCG	TCG	TCA	ACG	ACA	GGC	TCT	GGG	TCT	GGT	TCT	GCG	TCT	AGG	TCA	GCT	TCT	GCT	AAG	CAC
		63			72			81			90			99			108		
G	S	A	G	S	S	T	F	D	H	E	Q	H	Q	Q	H	Q	Q	Q	Q
GGA	AGT	GCC	GGT	TCC	AGT	ACC	TTT	GAT	CAT	GAA	CAA	CAT	CAA	CAA	CAT	CAA	CAA	CAA	CAA
		123			132			141			150			159			168		
Q	Q	K	R	Q	R	S	Q	S	E	A	R	Q	Q	Q	Q	Q	Q	Q	Q
CAA	CAA	AAG	CGC	CAG	CGG	TCG	CAA	TCA	GAA	GCA	CGA	CAA	CAG	CAG	CAG	CAA	CAG	CAA	CAG
		183			192			201			210			219			228		
Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	A	Q	H	H	A	H	S	T	Y	A
CAA	CAG	CAA	CAG	CAA	CAA	CAA	CAA	CAA	CAA	GCA	CAG	CAC	CAT	GCA	CAT	TCT	ACA	TAT	GCA
		243			252			261			270			279			288		
Q	R	P	Q	P	T	P	Q	Q	R	P	P	Q	N	L	L	T	P	A	S
CAA	AGA	CCC	CAA	CCC	ACC	CCC	CAA	CAA	CGA	CCA	CCC	CAA	AAC	CTA	CTG	ACA	CCT	GCT	TCA
		303			312			321			330			339			348		
T	T	G	A	S	V	G	P	L	Q	R	A	Y	S	V	S	L	A	A	R
ACC	ACT	GGT	GCC	AGC	GTC	GGC	CCG	CTC	CAA	CGC	GCA	TAC	TCG	GTT	TCA	TTA	GCT	GCG	AGA
		363			372			381			390			399			408		
Q	S	P	S	T	N	L	V	R	P	K	T	D	S	P	A	P	H	T	L
CAG	TCC	CCC	TCG	ACA	AAC	TTG	GTC	CGT	CCA	AAG	ACC	GAC	TCG	CCA	GCT	CCC	CAC	ACT	TTA
		423			432			441			450			459			468		
H	L	K	N	K	K	N	L	R	H	P	A	P	T	P	D	S	P	I	V
CAC	CTC	AAG	AAC	AAG	AAG	AAC	CTC	CGT	CAC	CCC	GCC	CCC	ACG	CCC	GAC	AGT	CCG	ATC	GTA
		483			492			501			510			519			528		
D	D	D	I	F	S	D	A	V	D	L	T	E	E	L	D	H	D	H	D
GAC	GAC	GAT	ATT	TTC	TCC	GAC	GCC	GTC	GAT	CTT	ACC	GAA	GAA	CTC	GAT	CAT	GAC	CAT	GAT
		543			552			561			570			579			588		
L	N	G	K	D	K	D	N	T	D	N	D	N	T	V	A	S	S	S	L
CTC	AAC	GGC	AAA	GAC	AAA	GAC	AAC	ACC	GAC	AAC	GAC	AAC	ACA	GTC	GCT	TCC	AGT	TCG	CTA
		603			612			621			630			639			648		
I	G	F	G	D	D	K	L	L	W	R	E	D	F	A	E	R	A	E	P
ATA	GGG	TTC	GGC	GAT	GAC	AAG	TTA	CTG	TGG	CGA	GAG	GAC	TTT	GCT	GAG	CGT	GCA	GAG	CCC
		663			672			681			690			699			708		
E	H	E	R	G	G	S	R	P	R	Q	V	K	K	R	K	I	S	N	D
GAA	CAT	GAA	AGA	GGT	GGG	AGC	AGG	CCT	CGC	CAG	GTC	AAG	AAA	CGG	AAG	ATA	TCG	AAT	GAC
		723			732			741			750			759			768		
Y	I	M	K	D	E	D	V	S	L	F	D	D	D	G	E	E	D	E	F
TAC	ATT	ATG	AAG	GAT	GAG	GAT	GTC	TCG	CTT	TTT	GAT	GAT	GAT	GGC	GAG	GAG	GAC	GAG	TTT
		783			792			801			810			819			828		
M	D	I	N	E	L	V	Q	G	D	R	E	S	T	P	K	P	K	A	T
ATG	GAT	ATC	AAT	GAG	CTA	GTT	CAG	GGG	GAT	CGG	GAA	AGT	ACT	CCG	AAG	CCA	AAG	GCT	ACA
		843			852			861			870			879			888		

FIG. 5/1

```
S   R   S   V   S   T   R   L   P   P   T   V   S   L   Q   R   G   R   S   P
TCG AGG TCT GTC TCG ACG AGG CTG CCG CCT ACA GTA TCG CTG CAA CGG GGT CGG TCT CCT
        903         912         921         930         939         948

K   R   K   E   A   S   V   E   K   R   T   T   E   N   Q   Q   Q   A   D   R
AAG AGG AAG GAG GCT TCA GTT GAA AAG CGC ACA ACG GAA AAC CAG CAA CAG GCT GAC AGA
        963         972         981         990         999         1008

E   D   E   P   S   F   M   S   S   P   D   V   D   N   S   R   K   R   K   S
GAA GAC GAA CCG TCG TTT ATG TCA AGT CCA GAT GTC GAC AAC TCC CGC AAG CGA AAG TCT
        1023        1032        1041        1050        1059        1068

S   G   S   P   T   G   L   T   T   P   R   P   Q   Q   K   Q   T   E   E   V
TCT GGA TCG CCC ACA GGT TTA ACG ACG CCA AGA CCC CAG CAG AAG CAA ACG GAA GAG GTC
        1083        1092        1101        1110        1119        1128

P   G   T   T   T   A   K   K   P   R   R   S   E   V   M   D   S   E   D   E
CCA GGT ACG ACC ACC GCC AAG AAG CCA CGG CGC AGT GAA GTG ATG GAC TCG GAG GAC GAG
        1143        1152        1161        1170        1179        1188

A   F   T   P   L   S   A   G   S   L   P   G   S   A   E   F   F   R   S   G
GCA TTC ACT CCT CTT TCT GCT GGG TCG CTG CCT GGG AGT GCG GAG TTC TTC AGA AGC GGT
        1203        1212        1221        1230        1239        1248

G   T   T   T   R   E   L   G   L   D   E   D   T   V   M   D   T   P   S   R
GGG ACC ACC ACA CGG GAA TTG GGT TTG GAC GAA GAC ACG GTT ATG GAC ACG CCT AGT AGG
        1263        1272        1281        1290        1299        1308

P   P   V   E   S   T   L   P   T   L   E   S   V   E   S   R   P   P   P   L
CCA CCG GTC GAG TCC ACT TTG CCA ACT CTC GAG TCT GTG GAA AGT CGA CCA CCC CCC CTG
        1323        1332        1341        1350        1359        1368

P   P   M   D   L   P   S   Q   R   K   P   L   E   P   L   N   T   P   R   N
CCG CCC ATG GAT CTA CCA TCA CAG CGA AAA CCG CTA GAG CCG TTG AAC ACT CCG CGC AAC
        1383        1392        1401        1410        1419        1428

Q   L   L   E   S   V   E   R   P   T   Q   Q   P   S   V   G   P   S   F   A
CAG CTG CTT GAG TCG GTC GAA AGG CCA ACA CAG CAG CCG TCG GTG GGG CCG AGT TTT GCA
        1443        1452        1461        1470        1479        1488

Q   S   S   T   L   A   E   S   S   L   P   P   S   M   P   P   P   S   E   D
CAG AGT AGC ACA CTC GCC GAA AGC TCC CTG CCG CCG TCA ATG CCG CCG CCA AGT GAA GAC
        1503        1512        1521        1530        1539        1548

P   L   N   T   R   E   N   S   N   L   E   E   F   D   Y   K   L   Y   K   P
CCC CTC AAC ACC AGG GAG AAC AGC AAC CTT GAG GAG TTC GAC TAC AAG CTT TAC AAA CCC
        1563        1572        1581        1590        1599        1608

L   L   D   L   F   V   N   A   P   A   I   L   E   R   E   L   S   A   V   N
CTG CTA GAT CTT TTC GTC AAC GCA CCC GCA ATC TTG GAA AGA GAA CTG AGC GCC GTT AAT
        1623        1632        1641        1650        1659        1668

D   E   L   Q   E   N   M   I   K   L   R   D   C   L   R   L   P   R   E   E
GAC GAG CTT CAG GAG AAC ATG ATC AAG CTG CGG GAC TGT CTG CGC CTG CCC AGG GAA GAA
        1683        1692        1701        1710        1719        1728

R   D   R   A   R   E   E   V   K   K   E   K   E   M   L   K   R   R   D   I
AGA GAC AGG GCA CGC GAA GAG GTG AAG AAG GAA AAG GAA ATG CTC AAG CGA CGG GAC ATT
        1743        1752        1761        1770        1779        1788

A   L   R   A   L   Q   D   E   H   K   L   Y   V   K   K   R   K   E   H   N
GCG CTC AGA GCC CTC CAG GAC GAA CAC AAG TTG TAC GTC AAG AAA CGC AAA GAG CAT AAT
        1803        1812        1821        1830        1839        1848
```

FIG. 5/2

```
L   I   N   E   E   I   V   R   A   Y   A   E   E   D   D   E   Y   E   D   Q
TTG ATC AAC GAG GAA ATC GTT CGC GCT TAT GCT GAA GAA GAC GAT GAG TAC GAG GAT CAG
        1863        1872        1881        1890        1899        1908

L   M   A   Q   L   D   K   L   D   D   E   V   E   A   I   V   K   S   L   T
TTA ATG GCG CAG CTG GAC AAG TTG GAT GAT GAG GTT GAG GCT ATC GTA AAG AGT CTG ACA
        1923        1932        1941        1950        1959        1968

R   L   I   V   A   A   G   I   T   E   K   S   F   D   L   K   K   E   E   E
AGG CTT ATT GTG GCG GCG GGG ATC ACG GAG AAG AGC TTT GAC CTA AAG AAG GAG GAG GAA
        1983        1992        2001        2010        2019        2028

E   E   E   E   K   P   I   I   I   A   T   P   T   P   S   T   R   T   E   A
GAG GAG GAG GAG AAG CCG ATC ATC ATA GCG ACT CCG ACA CCT TCG ACG AGG ACC GAG GCC
        2043        2052        2061        2070        2079        2088

P   V   L   P   T   T   E   Y   H   N   S   Q   Q   V   I   L   Q   T   Q   H
CCG GTT CTG CCG ACG ACC GAG TAT CAT AAT TCC CAG CAG GTC ATA TTG CAG ACT CAA CAT
        2103        2112        2121        2130        2139        2148

P   A   A   Q   Q   V   S   H   R   V   P   P   P   P   T   P   S   F   Q   T
CCT GCT GCG CAG CAG GTT TCT CAC CGG GTG CCA CCA CCT CCG ACA CCG AGT TTT CAA ACG
        2163        2172        2181        2190        2199        2208

A   R   Q   T   P   V   S   Y   Q   S   R   P   T   N   N   S   F   P   D   I
GCG CGC CAG ACT CCG GTG TCA TAT CAG AGC AGA CCG ACC AAC AAC TCC TTT CCT GAT ATC
        2223        2232        2241        2250        2259        2268

S   A   E   E   A   M   M   F   D   K   E   D   P   F   M   E   Q   Q   H   A
TCG GCG GAA GAA GCC ATG ATG TTC GAT AAA GAA GAC CCC TTC ATG GAA CAA CAG CAC GCC
        2283        2292        2301        2310        2319        2328

P   A   S   A   P   F   Q   A   T   L   P   Q   R   N   S   P   F   K   T   A
CCG GCC TCT GCT CCC TTC CAG GCC ACC CTT CCC CAG CGC AAC AGC CCT TTC AAA ACC GCC
        2343        2352        2361        2370        2379        2388

P   F   K   P   V   H   G   H   D   Y   F   D   D   E   D   D   D   A   D   L
CCG TTC AAG CCA GTC CAC GGC CAC GAT TAC TTT GAC GAT GAA GAC GAC GAT GCC GAC CTC
        2403        2412        2421        2430        2439        2448

L   A   A   V   D   S   A   E   T   Y   T   S   T   A   A   T   T   T   T   N
CTG GCA GCA GTA GAC AGC GCC GAG ACG TAT ACT TCT ACG GCC GCC ACC ACC ACC ACC AAC
        2463        2472        2481        2490        2499        2508

N   N   N   H   L   R   S   Q   S   V   M   S   T   S   T   A   T   T   I   K
AAC AAC AAT CAC TTA CGA TCA CAA TCG GTG ATG TCA ACA TCC ACG GCG ACC ACG ATC AAA
        2523        2532        2541        2550        2559        2568

P   R   K   R   N   E   N   A   N   A   K   K   P   K   S   V   H   A   K   L
CCG AGG AAA CGC AAC GAA AAT GCC AAT GCC AAG AAG CCC AAG TCC GTA CAT GCA AAG TTA
        2583        2592        2601        2610        2619        2628

S   M   P   P   E   K   M   K   Y   A   W   S   N   D   V   R   K   A   L   K
TCG ATG CCG CCC GAA AAG ATG AAG TAT GCG TGG TCG AAT GAT GTG AGG AAG GCT CTC AAG
        2643        2652        2661        2670        2679        2688

D   R   F   R   M   S   G   F   R   Q   N   Q   L   E   A   I   N   A   T   L
GAT AGG TTT CGG ATG TCG GGG TTC AGA CAG AAT CAG TTG GAG GCT ATT AAT GCT ACT TTG
        2703        2712        2721        2730        2739        2748

G   G   K   D   A   F   V   L   M   P   T   G   G   G   K   S   L   C   Y   Q
GGT GGT AAG GAT GCC TTT GTG TTG ATG CCG ACT GGT GGT GGA AAG TCT CTG TGC TAT CAG
        2763        2772        2781        2790        2799        2808
```

FIG. 5/3

```
 L   P   A   V   V   R   S   G   K   T   R   G   I   T   V   V   I   S   P   L
TTG CCG GCT GTA GTC AGG AGC GGC AAG ACG CGT GGT ATC ACA GTC GTC ATC TCC CCT CTG
        2823        2832        2841        2850        2859        2868

 L   S   L   M   L   D   Q   V   N   H   L   A   N   L   M   I   Q   A   Y   A
CTA AGT CTG ATG CTG GAT CAA GTC AAC CAT TTG GCA AAC CTG ATG ATC CAA GCT TAC GCT
        2883        2892        2901        2910        2919        2928

 F   N   G   D   M   N   S   E   M   R   R   M   V   F   Q   K   L   D   A   E
TTC AAC GGA GAC ATG AAC TCA GAA ATG CGC CGA ATG GTG TTT CAG AAG CTT GAT GCT GAG
        2943        2952        2961        2970        2979        2988

 H   P   E   H   E   L   Q   L   L   Y   V   T   P   E   M   V   S   K   N   Q
CAT CCT GAG CAT GAG CTC CAA CTG CTC TAT GTC ACC CCG GAG ATG GTG AGC AAG AAC CAG
        3003        3012        3021        3030        3039        3048

 T   F   V   N   K   M   M   D   L   Y   R   R   K   K   L   A   R   I   V   I
ACA TTC GTC AAC AAG ATG ATG GAC CTC TAC CGG AGG AAA AAG CTG GCT AGA ATT GTT ATC
        3063        3072        3081        3090        3099        3108

 D   E   A   H   C   V   S   Q   W   G   H   D   F   R   P   D   Y   K   A   I
GAC GAG GCT CAC TGC GTC AGT CAA TGG GGC CAT GAC TTC CGA CCC GAT TAC AAA GCT ATA
        3123        3132        3141        3150        3159        3168

 G   E   F   R   K   R   F   P   G   V   P   V   M   A   L   T   A   T   A   T
GGA GAG TTT CGT AAG AGG TTT CCC GGA GTT CCG GTC ATG GCT TTG ACA GCG ACG GCA ACA
        3183        3192        3201        3210        3219        3228

 Q   N   V   I   L   D   V   K   H   N   L   A   M   E   D   C   Q   T   F   S
CAG AAC GTC ATC TTG GAT GTC AAG CAT AAC CTG GCA ATG GAG GAC TGC CAG ACT TTC TCC
        3243        3252        3261        3270        3279        3288

 Q   S   F   N   R   P   N   L   Y   Y   E   V   R   M   K   E   Q   N   L   I
CAG AGC TTT AAT CGG CCG AAC CTC TAC TAT GAG GTC AGG ATG AAG GAG CAG AAT CTG ATT
        3303        3312        3321        3330        3339        3348

 A   R   I   A   E   L   I   K   E   K   Y   D   G   Q   T   G   I   I   Y   T
GCC CGC ATC GCA GAG TTG ATC AAG GAG AAG TAT GAC GGC CAG ACG GGT ATC ATC TAC ACA
        3363        3372        3381        3390        3399        3408

 L   S   R   K   S   A   E   N   I   A   K   N   L   Q   E   K   H   R   I   K
TTA TCA AGA AAG AGT GCC GAG AAC ATC GCC AAA AAT CTC CAG GAA AAA CAC CGC ATC AAA
        3423        3432        3441        3450        3459        3468

 A   K   H   Y   H   A   S   I   T   T   D   E   K   I   S   V   Q   H   E   W
GCA AAG CAC TAC CAT GCC TCC ATC ACC ACC GAC GAA AAG ATC AGC GTT CAA CAT GAA TGG
        3483        3492        3501        3510        3519        3528

 Q   T   G   R   V   K   V   V   V   A   T   I   A   F   G   M   G   I   D   K
CAA ACC GGC CGA GTC AAA GTC GTG GTA GCC ACC ATT GCC TTC GGC ATG GGC ATC GAC AAG
        3543        3552        3561        3570        3579        3588

 P   D   V   R   F   V   I   H   Q   H   I   P   K   S   L   E   G   Y   Y   Q
CCT GAC GTC CGC TTT GTT ATC CAC CAG CAC ATC CCC AAG TCG CTC GAA GGT TAC TAT CAA
        3603        3612        3621        3630        3639        3648

 E   T   G   R   A   G   R   D   G   K   P   S   D   C   Y   L   Y   F   A   Y
GAA ACC GGC CGC GCC GGA CGT GAC GGC AAG CCA TCG GAC TGC TAC TTG TAC TTT GCC TAT
        3663        3672        3681        3690        3699        3708

 G   D   I   Q   S   L   R   R   M   I   A   D   G   E   G   D   Y   A   Q   K
GGC GAC ATT CAA TCC CTA CGT CGT ATG ATC GCC GAC GGC GAA GGT GAC TAC GCG CAA AAG
        3723        3732        3741        3750        3759        3768
```

FIG 5/4

E R Q L Q M L N R V V S Y C E S Q H T C
GAG CGT CAG CTA CAA ATG CTC AAC CGT GTG GTC AGC TAT TGC GAG TCG CAG CAC ACG TGC
3783 3792 3801 3810 3819 3828

R R E E V L R Y F G E E F D Y R K C R D
CGG CGC GAA GAA GTG CTC CGC TAC TTT GGC GAG GAG TTT GAC TAC CGG AAG TGT AGA GAC
3843 3852 3861 3870 3879 3888

G C D N C R N G R I S K S T E M R D F T
GGA TGC GAT AAC TGC CGG AAC GGA CGC ATC TCG AAG TCG ACG GAG ATG AGG GAT TTT ACG
3903 3912 3921 3930 3939 3948

E I A F A A I E V V K S Q Q P I T L G K
GAA ATC GCC TTC GCC GCG ATC GAG GTG GTG AAG AGC CAG CAG CCC ATC ACG CTG GGC AAG
3963 3972 3981 3990 3999 4008

L C D I L M G K R K N E H G G V C H F G
CTG TGC GAC ATC CTG ATG GGC AAG AGA AAG AAC GAG CAC GGT GGC GTG TGT CAC TTT GGT
4023 4032 4041 4050 4059 4068

I A K G S T Q R E L Q R I V L Q L N F H
ATC GCC AAG GGG AGC ACG CAG AGG GAG CTG CAG AGG ATC GTG CTG CAG CTG AAT TTC CAC
4083 4092 4101 4110 4119 4128

K A L G E D N I M N G A G M P I T Y Y I
AAG GCG CTG GGC GAG GAC AAT ATC ATG AAT GGG GCG GGG ATG CCT ATT ACC TAC TAT ATT
4143 4152 4161 4170 4179 4188

T G P E A G A Y L Y N G K R L M L P V P
ACC GGC CCT GAA GCT GGT GCT TAC CTC TAC AAT GGC AAG CGG TTG ATG CTG CCA GTT CCC
4203 4212 4221 4230 4239 4248

S N K S V E P P S R S K Q R S R R V D E
TCA AAC AAG TCC GTC GAA CCC CCG TCT CGG TCT AAG CAG CGG AGC CGT CGA GTC GAT GAG
4263 4272 4281 4290 4299 4308

D M D E Q E L S T L Q R P P T S T N V S
GAT ATG GAT GAG CAA GAA CTT TCC ACC CTG CAA CGA CCG CCA ACA TCA ACA AAT GTC TCT
4323 4332 4341 4350 4359 4368

S P V R A T K K R S S K K A L P T L I A
TCA CCC GTT CGA GCC ACC AAG AAA CGA AGT TCC AAA AAG GCT TTA CCG ACC CTC ATC GCC
4383 4392 4401 4410 4419 4428

D Y E E P S S D G P H G P L H A N G Y E
GAC TAC GAA GAG CCC AGC TCC GAC GGT CCT CAC GGT CCT CTC CAC GCC AAC GGC TAT GAG
4443 4452 4461 4470 4479 4488

R D N F V V S D N V E P E E E E D A F E
CGC GAT AAC TTT GTC GTA TCC GAT AAT GTT GAA CCC GAA GAG GAA GAA GAT GCC TTC GAA
4503 4512 4521 4530 4539 4548

P V R P S R R G P S S R A T R P Q H R Q
CCT GTC CGC CCC TCG CGG CGC GGC CCA TCT TCT CGC GCT ACC CGC CCT CAA CAC CGC CAG
4563 4572 4581 4590 4599 4608

T T L Y D T L S H T Q Q S Q T V S Q H L
ACC ACC CTT TAT GAC ACC CTC TCC CAC ACC CAA CAA TCC CAA ACC GTC TCC CAA CAC CTC
4623 4632 4641 4650 4659 4668

A T L G P P I D A R T M H N P R Y A Q L
GCC ACT TTG GGT CCG CCC ATC GAC GCC CGC ACC ATG CAT AAC CCC CGC TAC GCC CAG CTT
4683 4692 4701 4710 4719 4728

FIG. 5/5

```
D   E   V   H   Q   D   I   V   D   A   F   V   E   E   V   K   V   F   E   E
GAC GAG GTC CAC CAG GAT ATT GTC GAT GCC TTT GTT GAA GAA GTC AAG GTC TTC GAG GAG
        4743        4752        4761        4770        4779        4788

D   F   R   N   R   N   H   M   R   K   P   I   F   T   E   T   Q   Y   R   E
GAC TTT CGC AAC AGG AAC CAC ATG CGC AAA CCC ATC TTT ACC GAG ACG CAG TAC CGT GAG
        4803        4812        4821        4830        4839        4848

M   A   I   R   W   T   R   S   L   D   A   M   R   A   I   P   D   I   N   Q
ATG GCA ATC CGG TGG ACG CGG TCG TTA GAC GCG ATG CGC GCG ATC CCG GAT ATC AAC CAG
        4863        4872        4881        4890        4899        4908

D   K   V   D   R   Y   G   A   K   F   I   P   L   V   E   R   F   W   G   N
GAT AAA GTA GAT CGG TAT GGT GCC AAA TTC ATC CCA CTT GTG GAG CGG TTC TGG GGG AAT
        4923        4932        4941        4950        4959        4968

Y   Q   E   M   M   G   G   G   Y   D   N   P   A   V   A   G   D   E   D   D
TAT CAG GAG ATG ATG GGG GGA GGG TAT GAT AAT CCT GCT GTG GCT GGC GAT GAG GAT GAT
        4983        4992        5001        5010        5019        5028

D   E   G   P   R   R   T   G   N   G   K   G   G   N   K   K   G   G   G   G
GAT GAG GGC CCC AGG AGG ACA GGA AAT GGA AAA GGG GGG AAT AAG AAG GGA GGA GGA GGA
        5043        5052        5061        5070        5079        5088

G   G   G   N   E   V   V   D   L   I   S   S   D   E   D   E   P   P   A   R
GGA GGA GGA AAT GAA GTA GTG GAT TTG ATT AGT AGT GAT GAG GAT GAA CCT CCG GCT CGT
        5103        5112        5121        5130        5139        5148

A   P   S   R   N   A   G   R   G   K   A   Q   S   T   R   G   G   Q   I   Q
GCA CCA TCG CGG AAT GCG GGG CGA GGA AAG GCA CAG TCG ACA CGT GGG GGA CAA ATC CAA
        5163        5172        5181        5190        5199        5208

D   K   G   R   A   V   N   R   R   G   E   P   I   A   E   E   D   E   E   D
GAT AAA GGC CGA GCA GTC AAC CGC CGC GGA GAA CCC ATC GCC GAA GAA GAC GAA GAA GAC
        5223        5232        5241        5250        5259        5268

Y   G   L   S   D   P   D   I   D   A   I   D   P   D   A   I   T   A   S   D
TAC GGG CTA AGC GAC CCC GAT ATC GAC GCC ATC GAT CCA GAC GCC ATC ACC GCC TCC GAC
        5283        5292        5301        5310        5319        5328

N   S   D   E   E   D   D   D   D   D   D   E   D   L   E   S   S   R   Y   F
AAC TCC GAC GAA GAA GAT GAT GAT GAT GAT GAC GAA GAC CTC GAA TCC TCC CGC TAC TTC
        5343        5352        5361        5370        5379        5388

S   G   S   T   G   P   P   V   S   K   A   V   Q   D   A   R   L   R   E   Q
TCC GGC TCA ACA GGC CCG CCC GTC TCC AAA GCC GTG CAG GAT GCT CGA CTC CGT GAA CAA
        5403        5412        5421        5430        5439        5448

L   S   M   Y   A   S   G   G   S   S   S   K   G   S   Y   G   S   G   R   A
CTT TCC ATG TAC GCC TCC GGC GGC AGC TCT TCG AAA GGT AGC TAC GGC TCA GGG CGC GCA
        5463        5472        5481        5490        5499        5508

S   G   G   S   S   S   R   A   S   G   S   G   W   R   G   G   G   A   G   G
TCA GGA GGA TCT TCG TCG AGA GCG TCG GGA TCA GGA TGG AGA GGT GGA GGA GCA GGT GGG
        5523        5532        5541        5550        5559        5568

K   K   Y   Y   R   K   K   R   A   G   S   S   A   A   G   G   G   G   A   G
AAG AAA TAC TAC AGG AAG AAG AGG GCT GGT TCT TCG GCT GCT GGT GGT GGT GGT GCA GGA
        5583        5592        5601        5610        5619        5628

G   G   G   V   T   K   R   K   A   S   G   S   G   A   K   T   A   R   K   R
GGA GGG GGA GTT ACA AAA CGG AAG GCG AGT GGG AGT GGC GCG AAG ACG GCG AGG AAG AGG
        5643        5652        5661        5670        5679        5688
```

FIG. 5/6

```
G   A   S   T   A   P   K   T   T   T   R   G   G   G   S   G   A   G   S   R
GGT GCA TCT ACT GCG CCG AAG ACA ACG ACG AGA GGG GGA GGA TCT GGA GCT GGG TCT AGA
        5703        5712        5721        5730        5739        5748

G   G   G   A   G   G   A   G   G   A   G   A   G   A   G   A   G   G   G   K
GGA GGC GGT GCT GGT GGT GCT GGT GGT GCT GGT GCT GGT GCT GGT GCT GGA GGA GGG AAA
        5763        5772        5781        5790        5799        5808

R   G   G   G   G   G   G   G   M   G   G   I   S   V   M   P   H   .
AGG GGT GGT GGA GGT GGA GGA GGA ATG GGA GGG ATA AGT GTT ATG CCT CAT TAG
        5823        5832        5841        5850        5859
```

FIG. 5/7

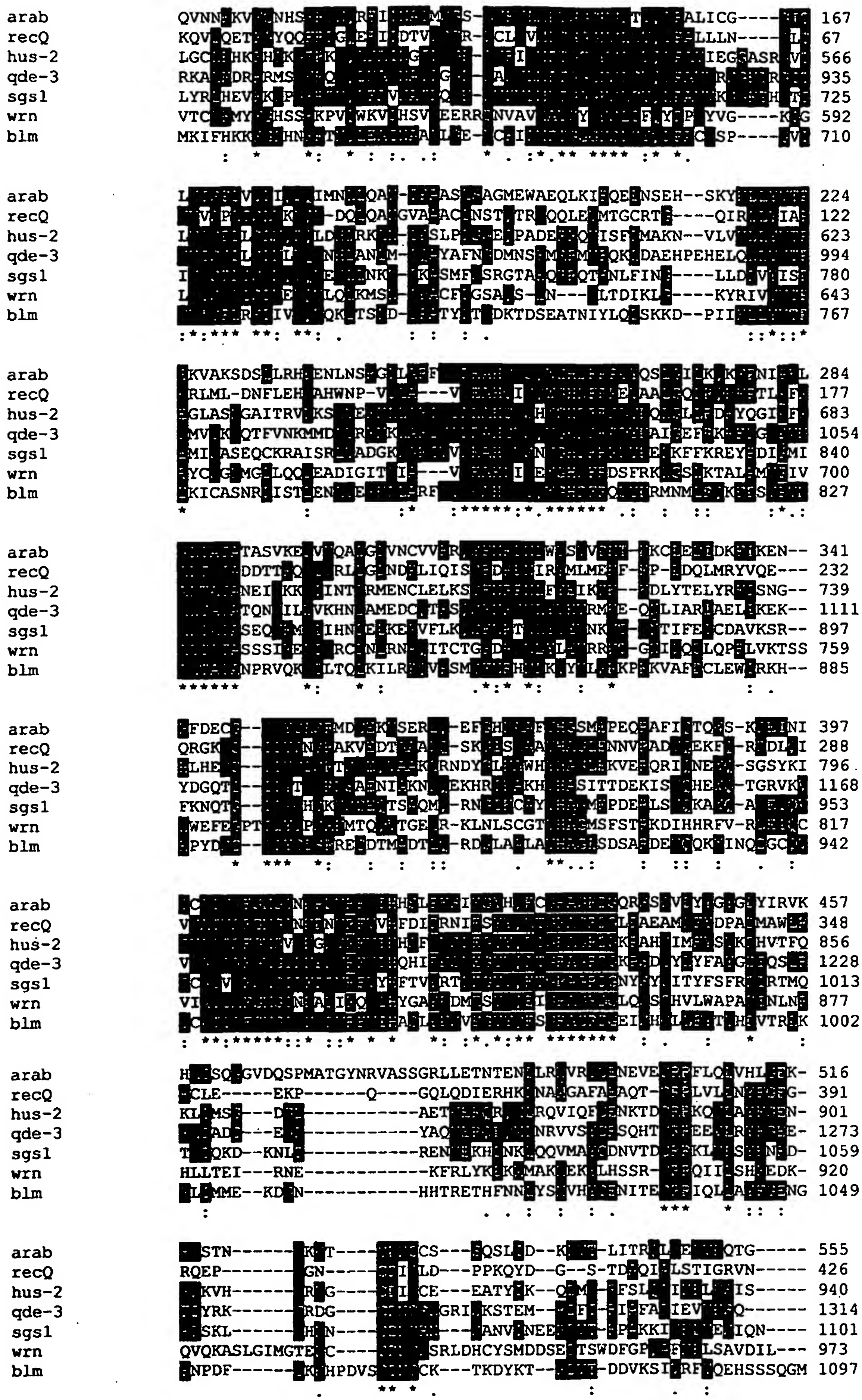


FIG. 6

SEQUENCE LISTING

<110> Università degli studi di Roma La Sapienza
Macino, Giuseppe
Cogoni, Carlo

<120> Isolation and characterization of a N. crassa silencing gene and uses therof

<130> PC

<140>
<141>

<160> 3

<170> PatentIn Ver. 2.1

<210> 1
<211> 6916
<212> DNA
<213> Neurospora crassa

<400> 1
ccaccaccac caccaccacc acaacaacca actcaacaac gaactcaaca accaacccca 60
aacccaacct cgacctcaac ctcaacctca acccttgcga cctcgagata cacaaacaca 120
tcctcgacaa atgacgcccg acccgctaca cgccaacaga ttgcccccga ggtaggagca 180
tcgacgcatc aggattcagt tggacttgga gaaggaggag gaggaggaat ggcgaagctc 240
tcagtcaaga acaacctgcc acggccgcac ttggtctcct tgtcgtcgtc aacgacaggc 300
tctgggtctg gttctgcgtc taggtcagct tctgctaagc acggaagtgc cggttccagt 360
acctttgatc atgaacaaca tcaacaacat caacaacaac aacaacaaaa gcgccagcgg 420
tcgcaatcag aagcacgaca acagcagcag caacagcaac agcaacagca acagcaacaa 480
caacaacaac aagcacagca ccatgcacat tctacatatg cacaaagacc ccaacccacc 540
ccccaacaac gaccacccca aaacctactg acacctgctt caaccactgg tgccagcgtc 600
ggcccgctcc aacgcgcata ctcggtttca ttagctgcga gacagtcccc ctcgacaaac 660
ttggtccgtc caaagaccga ctcgccagct ccccacactt tacacctcaa gaacaagaag 720
aacctccgtc accccgcccc cacgcccgac agtccgatcg tagacgacga tatttctcc 780
gacgccgtcg atcttaccga agaactcgat catgaccatg atctcaacgg caaagacaaa 840
gacaacaccg acaacgacaa cacagtcgct tccagttcgc taatagggtt cggcgatgac 900
aagttactgt ggcgagagga ctttgctgag cgtgcagagc ccgaacatga aagaggtggg 960
agcaggcctc gccaggtcaa gaaacggaag atatcgaatg actacattat gaaggatgag 1020
gatgtctcgc tttttgatga tgatggcgag gaggacgagt ttatggatat caatgagcta 1080
gttcaggggg atcgggaaag tactccgaag ccaaaggcta catcgaggtc tgtctcgacg 1140
aggctgccgc ctacagtatc gctgcaacgg ggtcggtctc ctaagaggaa ggaggcttca 1200
gttgaaaagc gcacaacgga aaaccagcaa caggctgaca gagaagacga accgtcgttt 1260
atgtcaagtc cagatgtcga caactcccgc aagcgaaagt cttctggatc gcccacaggt 1320
ttaacgacgc caagacccca gcagaagcaa acggaagagg tcccaggtac gaccaccgcc 1380
aagaagccac ggcgcagtga agtgatggac tcggaggacg aggcattcac tcctctttct 1440

gctgggtcgc tgcctgggag tgcggagttc ttcagaagcg gtgggaccac cacacgggaa 1500
ttgggtttgg acgaagacac ggttatggac acgcctagta ggccaccggt cgagtccact 1560
ttgccaactc tcgagtctgt ggaaagtcga ccaccccccc tgccgcccat ggatctacca 1620
tcacagcgaa aaccgctaga gccgttgaac actccgcgca accagctgct tgagtcggtc 1680
gaaaggccaa cacagcagcc gtcggtgggg ccgagttttg cacagagtag cacactcgcc 1740
gaaagctccc tgccgccgtc aatgccgccg ccaagtgaag accccctcaa caccagggag 1800
aacagcaacc ttgaggagtt cgactacaag ctttacaaac ccctgctaga tcttttcgtc 1860
aacgcacccg caatcttgga aagagaactg agcgccgtta atgacgagct tcaggagaac 1920
atgatcaagc tgcgggactg tctgcgcctg cccagggaag aaagagacag ggcacgcgaa 1980
gaggtgaaga aggaaaagga aatgctcaag cgacgggaca ttgcgctcag agccctccag 2040
gacgaacaca agttgtacgt caagaaacgc aaagagcata atttgatcaa cgaggaaatc 2100
gttcgcgctt atgctgaaga agacgatgag tacgaggatc agttaatggc gcagctggac 2160
aagttggatg atgaggttga ggctatcgta aagagtctga caaggcttat tgtggcggcg 2220
gggatcacgg agaagagctt tgacctaaag aaggaggagg aagaggagga ggagaagccg 2280
atcatcatag cgactccgac accttcgacg aggaccgagg ccccggttct gccgacgacc 2340
gagtatcata attcccagca ggtcatattg cagactcaac atcctgctgc gcagcaggtt 2400
tctcaccggg tgccaccacc tccgacaccg agttttcaaa cggcgcgcca gactccggtg 2460
tcatatcaga gcagaccgac caacaactcc tttcctgata tctcggcgga agaagccatg 2520
atgttcgata aagaagaccc cttcatggaa caacagcacg ccccggcctc tgctcccttc 2580
caggccaccc ttccccagcg caacagccct ttcaaaaccg ccccgttcaa gccagtccac 2640
ggccacgatt actttgacga tgaagacgac gatgccgacc tcctggcagc agtagacagc 2700
gccgagacgt atacttctac ggccgccacc accaccacca acaacaacaa tcacttacga 2760
tcacaatcgg tgatgtcaac atccacggcg accacgatca aaccgaggaa acgcaacgaa 2820
aatgccaatg ccaagaagcc caagtccgta catgcaaagt tatcgatgcc gcccgaaaag 2880
atgaagtatg cgtggtcgaa tgatgtgagg aaggctctca aggataggtt tcggatgtcg 2940
gggttcagac agaatcagtt ggaggctatt aatgctactt tgggtggtaa ggtgagttct 3000
ctgtccttta cctatctggg agagaccaag aaggagagag agagagagag gaggggaaga 3060
cgaaaatgga ctttgctgac tctagaaagg atgcctttgt gttgatgccg actggtggtg 3120
gaaagtctct gtgctatcag ttgccggctg tagtcaggag cggcaagacg cgtggtatca 3180
cagtcgtcat ctcccctctg ctaagtctga tgctggatca agtcaaccat ttggcaaacc 3240
tgatgatcca agcttacgct ttcaacggag acatgaactc agaaatgcgc cgaatggtgt 3300
ttcagaagct tgatgctgag catcctgagc atgagctcca actgctctat gtcaccccgg 3360
agatggtgag caagaaccag acattcgtca acaagatgat ggacctctac cggaggaaaa 3420
agctggctag aattgttatc gacgaggctc actgcgtcag tcaatggggc catgacttcc 3480
gacccgatta caaagctata ggagagtttc gtaagaggtt tcccggagtt ccggtcatgg 3540
ctttgacagc gacggcaaca cagaacgtca tcttggatgt caagcataac ctggcaatgg 3600
aggactgcca gactttctcc cagagcttta atcggccgaa cctctactat gaggtcagga 3660
tgaaggagca gaatctgatt gcccgcatcg cagagttgat caaggagaag tatgacggcc 3720
agacgggtat catctacaca ttatcaagaa agagtgccga gaacatcgcc aaaaatctcc 3780
aggaaaaaca ccgcatcaaa gcaaagcact accatgcctc catcaccacc gacgaaaaga 3840
tcagcgttca acatgaatgg caaaccggcc gagtcaaagt cgtggtagcc accattgcct 3900
tcggcatggg catcgacaag cctgacgtcc gctttgttat ccaccagcac atccccaagt 3960
cgctcgaagg ttactatcaa gaaaccggcc gcgccggacg tgacggcaag ccatcggact 4020
gctacttgta ctttgcctat ggcgacattc aatccctacg tcgtatgatc gccgacggcg 4080
aaggtgacta cgcgcaaaag gagcgtcagc tacaaatgct caaccgtgtg gtcagctatt 4140
gcgagtcgca gcacacgtgc cggcgcgaag aagtgctccg ctactttggc gaggagtttg 4200
actaccggaa gtgtagagac ggatgcgata actgccggaa cggacgcatc tcgaagtcga 4260
cggagatgag ggattttacg gaaatcgcct tcgccgcgat cgaggtggtg aagagccagc 4320

agcccatcac gctgggcaag ctgtgcgaca tcctgatggg caagagaaag aacgagcacg 4380
gtggcgtgtg tcactttggt atcgccaagg ggagcacgca gagggagctg cagaggatcg 4440
tgctgcagct gaatttccac aaggcgctgg gcgaggacaa tatcatgaat ggggcgggga 4500
tgcctattac ctactatatt gtgagtgctg tcccggttgg tcttgcatat ctggctttgt 4560
tgctttgcta acacagcagc tcgtacagac cggccctgaa gctggtgctt acctctacaa 4620
tggcaagcgg ttgatgctgc cagttccctc aaacaagtcc gtcgaacccc cgtctcggtc 4680
taagcagcgg agccgtcgag tcgatgagga tatggatgag caagaacttt ccaccctgca 4740
acgaccgcca acatcaacaa atgtctcttc acccgttcga gccaccaaga aacgaagttc 4800
caaaaaggct ttaccgaccc tcatcgccga ctacgaagag cccagctccg acggtcctca 4860
cggtcctctc cacgccaacg gctatgagcg cgataacttt gtcgtatccg ataatgttga 4920
acccgaagag gaagaagatg ccttcgaacc tgtccgcccc tcgcggcgcg gcccatcttc 4980
tcgcgctacc cgccctcaac accgccagac caccctttat gacaccctct cccacaccca 5040
acaatcccaa accgtctccc aacacctcgc cactttgggt ccgcccatcg acgcccgcac 5100
catgcataac ccccgctacg cccagcttga cgaggtccac caggatattg tcgatgcctt 5160
tgttgaagaa gtcaaggtct tcgaggagga ctttcgcaac aggaaccaca tgcgcaaacc 5220
catctttacc gagacgcagt accgtgagat ggcaatccgg tggacgcggt cgttagacgc 5280
gatgcgcgcg atcccggata tcaaccagga taaagtagat cggtatggtg ccaaattcat 5340
cccacttgtg gagcggttct gggggaatta tcaggagatg atggggggag ggtatgataa 5400
tcctgctgtg gctggcgatg aggatgatga tgagggcccc aggaggacag gaaatggaaa 5460
agggggggaat aagaagggag gaggaggagg aggaggaaat gaagtagtgg atttgattag 5520
tagtgatgag gatgaacctc cggctcgtgc accatcgcgg aatgcggggc gaggaaaggc 5580
acagtcgaca cgtgggggac aaatccaaga taaaggccga gcagtcaacc gccgcggaga 5640
acccatcgcc gaagaagacg aagaagacta cgggctaagc gaccccgata tcgacgccat 5700
cgatccagac gccatcaccg cctccgacaa ctccgacgaa gaagatgatg atgatgatga 5760
cgaagacctc gaatcctccc gctacttctc cggctcaaca ggcccgcccg tctccaaagc 5820
cgtgcaggat gctcgactcc gtgaacaact ttccatgtac gcctccggcg gcagctcttc 5880
gaaaggtagc tacggctcag ggcgcgcatc aggaggatct tcgtcgagag cgtcgggatc 5940
aggatggaga ggtggaggag caggtgggaa gaaatactac aggaagaaga gggctggttc 6000
ttcggctgct ggtggtggtg gtgcaggagg agggggagtt acaaaacgga aggcgagtgg 6060
gagtggcgcg aagacggcga ggaagagggg tgcatctact gcgccgaaga caacgacgag 6120
agggggagga tctggagctg ggtctagagg aggcggtgct ggtggtgctg gtggtgctgg 6180
tgctggtgct ggtgctggag gagggaaaag ggtggtgga ggtggaggag gaatgggagg 6240
gataagtgtt atgcctcatt agctatttta tagcatatcg catttataca gtgtcttatg 6300
gaagggagga ggagaagaag aaggataagc tggcataagc ttgaaccggc caggccaaaa 6360
tggccagaga gctcaccggg caatcgagct tgaaatgagc ttgacatatt aggtattccc 6420
gagaatatag cgggattaca aggcacttac tttaccaagt cgaaagggac gagccaaatc 6480
tatggtactc gccagttgcg caacgttgag ttttatcatt cgtggagttt tcatcgtgga 6540
gtttttatta tcaactattc gttgtatagt tttcgttgta gatgttagtt ccggacgatc 6600
aaaaggggaa gtgtggaaca gagaagtcga aaggacaagc caaaatgaca tggcagtgtc 6660
cagtcagata ccctccagac aaaaccagac accaataaca aaccttcaa ccataacacc 6720
agcaaagcca atccttaggt acctacctag ggtagggtag gtccaggaat gtcttcccca 6780
aaggtacctc tacttattca tgttacgctc catcagtccc atcgcttagc atcgctgccc 6840
ggttacctat ctctacctct acctctacct ctacctctac ctctacctct acctctatct 6900
ctacctctac ctctac 6916

<210> 2
<211> 5868

<212> DNA
<213> Neurospora crassa

<220>
<221> CDS
<222> (1)..(5868)

<400> 2
atg gcg aag ctc tca gtc aag aac aac ctg cca cgg ccg cac ttg gtc 48
Met Ala Lys Leu Ser Val Lys Asn Asn Leu Pro Arg Pro His Leu Val
1 5 10 15

tcc ttg tcg tcg tca acg aca ggc tct ggg tct ggt tct gcg tct agg 96
Ser Leu Ser Ser Ser Thr Thr Gly Ser Gly Ser Gly Ser Ala Ser Arg
20 25 30

tca gct tct gct aag cac gga agt gcc ggt tcc agt acc ttt gat cat 144
Ser Ala Ser Ala Lys His Gly Ser Ala Gly Ser Ser Thr Phe Asp His
35 40 45

gaa caa cat caa caa cat caa caa caa caa caa caa aag cgc cag cgg 192
Glu Gln His Gln Gln His Gln Gln Gln Gln Gln Gln Lys Arg Gln Arg
50 55 60

tcg caa tca gaa gca cga caa cag cag cag caa cag caa cag caa cag 240
Ser Gln Ser Glu Ala Arg Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
65 70 75 80

caa cag caa caa caa caa caa caa gca cag cac cat gca cat tct aca 288
Gln Gln Gln Gln Gln Gln Gln Gln Ala Gln His His Ala His Ser Thr
85 90 95

tat gca caa aga ccc caa ccc acc ccc caa caa cga cca ccc caa aac 336
Tyr Ala Gln Arg Pro Gln Pro Thr Pro Gln Gln Arg Pro Pro Gln Asn
100 105 110

cta ctg aca cct gct tca acc act ggt gcc agc gtc ggc ccg ctc caa 384
Leu Leu Thr Pro Ala Ser Thr Thr Gly Ala Ser Val Gly Pro Leu Gln
115 120 125

cgc gca tac tcg gtt tca tta gct gcg aga cag tcc ccc tcg aca aac 432
Arg Ala Tyr Ser Val Ser Leu Ala Ala Arg Gln Ser Pro Ser Thr Asn
130 135 140

ttg gtc cgt cca aag acc gac tcg cca gct ccc cac act tta cac ctc 480
Leu Val Arg Pro Lys Thr Asp Ser Pro Ala Pro His Thr Leu His Leu
145 150 155 160

aag aac aag aag aac ctc cgt cac ccc gcc ccc acg ccc gac agt ccg 528
Lys Asn Lys Lys Asn Leu Arg His Pro Ala Pro Thr Pro Asp Ser Pro
165 170 175

atc gta gac gac gat att ttc tcc gac gcc gtc gat ctt acc gaa gaa 576
Ile Val Asp Asp Asp Ile Phe Ser Asp Ala Val Asp Leu Thr Glu Glu
180 185 190

ctc gat cat gac cat gat ctc aac ggc aaa gac aaa gac aac acc gac 624
Leu Asp His Asp His Asp Leu Asn Gly Lys Asp Lys Asp Asn Thr Asp
195 200 205

aac gac aac aca gtc gct tcc agt tcg cta ata ggg ttc ggc gat gac 672
Asn Asp Asn Thr Val Ala Ser Ser Ser Leu Ile Gly Phe Gly Asp Asp
210 215 220

aag tta ctg tgg cga gag gac ttt gct gag cgt gca gag ccc gaa cat 720
Lys Leu Leu Trp Arg Glu Asp Phe Ala Glu Arg Ala Glu Pro Glu His
225 230 235 240

gaa aga ggt ggg agc agg cct cgc cag gtc aag aaa cgg aag ata tcg 768
Glu Arg Gly Gly Ser Arg Pro Arg Gln Val Lys Lys Arg Lys Ile Ser
245 250 255

aat gac tac att atg aag gat gag gat gtc tcg ctt ttt gat gat gat 816
Asn Asp Tyr Ile Met Lys Asp Glu Asp Val Ser Leu Phe Asp Asp Asp
260 265 270

ggc gag gag gac gag ttt atg gat atc aat gag cta gtt cag ggg gat 864
Gly Glu Glu Asp Glu Phe Met Asp Ile Asn Glu Leu Val Gln Gly Asp
275 280 285

cgg gaa agt act ccg aag cca aag gct aca tcg agg tct gtc tcg acg 912
Arg Glu Ser Thr Pro Lys Pro Lys Ala Thr Ser Arg Ser Val Ser Thr
290 295 300

agg ctg ccg cct aca gta tcg ctg caa cgg ggt cgg tct cct aag agg 960
Arg Leu Pro Pro Thr Val Ser Leu Gln Arg Gly Arg Ser Pro Lys Arg
305 310 315 320

aag gag gct tca gtt gaa aag cgc aca acg gaa aac cag caa cag gct 1008
Lys Glu Ala Ser Val Glu Lys Arg Thr Thr Glu Asn Gln Gln Gln Ala
325 330 335

gac aga gaa gac gaa ccg tcg ttt atg tca agt cca gat gtc gac aac 1056
Asp Arg Glu Asp Glu Pro Ser Phe Met Ser Ser Pro Asp Val Asp Asn
340 345 350

tcc cgc aag cga aag tct tct gga tcg ccc aca ggt tta acg acg cca 1104
Ser Arg Lys Arg Lys Ser Ser Gly Ser Pro Thr Gly Leu Thr Thr Pro
355 360 365

aga ccc cag cag aag caa acg gaa gag gtc cca ggt acg acc acc gcc 1152
Arg Pro Gln Gln Lys Gln Thr Glu Glu Val Pro Gly Thr Thr Thr Ala
370 375 380

aag aag cca cgg cgc agt gaa gtg atg gac tcg gag gac gag gca ttc 1200
Lys Lys Pro Arg Arg Ser Glu Val Met Asp Ser Glu Asp Glu Ala Phe
385 390 395 400

act cct ctt tct gct ggg tcg ctg cct ggg agt gcg gag ttc ttc aga 1248
Thr Pro Leu Ser Ala Gly Ser Leu Pro Gly Ser Ala Glu Phe Phe Arg
405 410 415

agc ggt ggg acc acc aca cgg gaa ttg ggt ttg gac gaa gac acg gtt 1296
Ser Gly Gly Thr Thr Thr Arg Glu Leu Gly Leu Asp Glu Asp Thr Val
420 425 430

atg gac acg cct agt agg cca ccg gtc gag tcc act ttg cca act ctc 1344
Met Asp Thr Pro Ser Arg Pro Pro Val Glu Ser Thr Leu Pro Thr Leu
435 440 445

gag tct gtg gaa agt cga cca ccc ccc ctg ccg ccc atg gat cta cca 1392
Glu Ser Val Glu Ser Arg Pro Pro Pro Leu Pro Pro Met Asp Leu Pro
450 455 460

tca cag cga aaa ccg cta gag ccg ttg aac act ccg cgc aac cag ctg 1440
Ser Gln Arg Lys Pro Leu Glu Pro Leu Asn Thr Pro Arg Asn Gln Leu
465 470 475 480

ctt gag tcg gtc gaa agg cca aca cag cag ccg tcg gtg ggg ccg agt 1488
Leu Glu Ser Val Glu Arg Pro Thr Gln Gln Pro Ser Val Gly Pro Ser
485 490 495

ttt gca cag agt agc aca ctc gcc gaa agc tcc ctg ccg ccg tca atg 1536
Phe Ala Gln Ser Ser Thr Leu Ala Glu Ser Ser Leu Pro Pro Ser Met
500 505 510

ccg ccg cca agt gaa gac ccc ctc aac acc agg gag aac agc aac ctt 1584
Pro Pro Pro Ser Glu Asp Pro Leu Asn Thr Arg Glu Asn Ser Asn Leu
515 520 525

gag gag ttc gac tac aag ctt tac aaa ccc ctg cta gat ctt ttc gtc 1632
Glu Glu Phe Asp Tyr Lys Leu Tyr Lys Pro Leu Leu Asp Leu Phe Val
530 535 540

aac gca ccc gca atc ttg gaa aga gaa ctg agc gcc gtt aat gac gag 1680
Asn Ala Pro Ala Ile Leu Glu Arg Glu Leu Ser Ala Val Asn Asp Glu
545 550 555 560

ctt cag gag aac atg atc aag ctg cgg gac tgt ctg cgc ctg ccc agg 1728
Leu Gln Glu Asn Met Ile Lys Leu Arg Asp Cys Leu Arg Leu Pro Arg
565 570 575

gaa gaa aga gac agg gca cgc gaa gag gtg aag aag gaa aag gaa atg 1776
Glu Glu Arg Asp Arg Ala Arg Glu Glu Val Lys Lys Glu Lys Glu Met
580 585 590

ctc aag cga cgg gac att gcg ctc aga gcc ctc cag gac gaa cac aag 1824
Leu Lys Arg Arg Asp Ile Ala Leu Arg Ala Leu Gln Asp Glu His Lys
595 600 605

ttg tac gtc aag aaa cgc aaa gag cat aat ttg atc aac gag gaa atc 1872
Leu Tyr Val Lys Lys Arg Lys Glu His Asn Leu Ile Asn Glu Glu Ile
610 615 620

gtt cgc gct tat gct gaa gaa gac gat gag tac gag gat cag tta atg 1920
Val Arg Ala Tyr Ala Glu Glu Asp Asp Glu Tyr Glu Asp Gln Leu Met
625 630 635 640

gcg cag ctg gac aag ttg gat gat gag gtt gag gct atc gta aag agt 1968
Ala Gln Leu Asp Lys Leu Asp Asp Glu Val Glu Ala Ile Val Lys Ser
645 650 655

ctg aca agg ctt att gtg gcg gcg ggg atc acg gag aag agc ttt gac 2016
Leu Thr Arg Leu Ile Val Ala Ala Gly Ile Thr Glu Lys Ser Phe Asp
660 665 670

cta aag aag gag gag gaa gag gag gag gag aag ccg atc atc ata gcg 2064
Leu Lys Lys Glu Glu Glu Glu Glu Glu Glu Lys Pro Ile Ile Ile Ala
675 680 685

act ccg aca cct tcg acg agg acc gag gcc ccg gtt ctg ccg acg acc 2112
Thr Pro Thr Pro Ser Thr Arg Thr Glu Ala Pro Val Leu Pro Thr Thr
690 695 700

gag tat cat aat tcc cag cag gtc ata ttg cag act caa cat cct gct 2160
Glu Tyr His Asn Ser Gln Gln Val Ile Leu Gln Thr Gln His Pro Ala
705 710 715 720

gcg cag cag gtt tct cac cgg gtg cca cca cct ccg aca ccg agt ttt 2208
Ala Gln Gln Val Ser His Arg Val Pro Pro Pro Pro Thr Pro Ser Phe
725 730 735

caa acg gcg cgc cag act ccg gtg tca tat cag agc aga ccg acc aac 2256
Gln Thr Ala Arg Gln Thr Pro Val Ser Tyr Gln Ser Arg Pro Thr Asn
740 745 750

aac tcc ttt cct gat atc tcg gcg gaa gaa gcc atg atg ttc gat aaa 2304
Asn Ser Phe Pro Asp Ile Ser Ala Glu Glu Ala Met Met Phe Asp Lys
755 760 765

gaa gac ccc ttc atg gaa caa cag cac gcc ccg gcc tct gct ccc ttc 2352
Glu Asp Pro Phe Met Glu Gln Gln His Ala Pro Ala Ser Ala Pro Phe
770 775 780

cag gcc acc ctt ccc cag cgc aac agc cct ttc aaa acc gcc ccg ttc 2400
Gln Ala Thr Leu Pro Gln Arg Asn Ser Pro Phe Lys Thr Ala Pro Phe
785 790 795 800

aag cca gtc cac ggc cac gat tac ttt gac gat gaa gac gac gat gcc 2448
Lys Pro Val His Gly His Asp Tyr Phe Asp Asp Glu Asp Asp Asp Ala
805 810 815

gac ctc ctg gca gca gta gac agc gcc gag acg tat act tct acg gcc 2496
Asp Leu Leu Ala Ala Val Asp Ser Ala Glu Thr Tyr Thr Ser Thr Ala
820 825 830

gcc acc acc acc acc aac aac aac aat cac tta cga tca caa tcg gtg 2544
Ala Thr Thr Thr Thr Asn Asn Asn Asn His Leu Arg Ser Gln Ser Val
835 840 845

atg tca aca tcc acg gcg acc acg atc aaa ccg agg aaa cgc aac gaa 2592
Met Ser Thr Ser Thr Ala Thr Thr Ile Lys Pro Arg Lys Arg Asn Glu
850 855 860

aat gcc aat gcc aag aag ccc aag tcc gta cat gca aag tta tcg atg 2640
Asn Ala Asn Ala Lys Lys Pro Lys Ser Val His Ala Lys Leu Ser Met
865 870 875 880

ccg ccc gaa aag atg aag tat gcg tgg tcg aat gat gtg agg aag gct 2688
Pro Pro Glu Lys Met Lys Tyr Ala Trp Ser Asn Asp Val Arg Lys Ala
885 890 895

ctc aag gat agg ttt cgg atg tcg ggg ttc aga cag aat cag ttg gag 2736
Leu Lys Asp Arg Phe Arg Met Ser Gly Phe Arg Gln Asn Gln Leu Glu
900 905 910

gct att aat gct act ttg ggt ggt aag gat gcc ttt gtg ttg atg ccg 2784
Ala Ile Asn Ala Thr Leu Gly Gly Lys Asp Ala Phe Val Leu Met Pro
915 920 925

act ggt ggt gga aag tct ctg tgc tat cag ttg ccg gct gta gtc agg 2832
Thr Gly Gly Gly Lys Ser Leu Cys Tyr Gln Leu Pro Ala Val Val Arg
930 935 940

agc ggc aag acg cgt ggt atc aca gtc gtc atc tcc cct ctg cta agt 2880
Ser Gly Lys Thr Arg Gly Ile Thr Val Val Ile Ser Pro Leu Leu Ser
945 950 955 960

ctg atg ctg gat caa gtc aac cat ttg gca aac ctg atg atc caa gct 2928
Leu Met Leu Asp Gln Val Asn His Leu Ala Asn Leu Met Ile Gln Ala
965 970 975

tac gct ttc aac gga gac atg aac tca gaa atg cgc cga atg gtg ttt 2976
Tyr Ala Phe Asn Gly Asp Met Asn Ser Glu Met Arg Arg Met Val Phe
980 985 990

cag aag ctt gat gct gag cat cct gag cat gag ctc caa ctg ctc tat 3024
Gln Lys Leu Asp Ala Glu His Pro Glu His Glu Leu Gln Leu Leu Tyr
995 1000 1005

gtc acc ccg gag atg gtg agc aag aac cag aca ttc gtc aac aag atg 3072
Val Thr Pro Glu Met Val Ser Lys Asn Gln Thr Phe Val Asn Lys Met
1010 1015 1020

atg gac ctc tac cgg agg aaa aag ctg gct aga att gtt atc gac gag 3120
Met Asp Leu Tyr Arg Arg Lys Lys Leu Ala Arg Ile Val Ile Asp Glu
1025 1030 1035 1040

gct cac tgc gtc agt caa tgg ggc cat gac ttc cga ccc gat tac aaa 3168
Ala His Cys Val Ser Gln Trp Gly His Asp Phe Arg Pro Asp Tyr Lys
1045 1050 1055

gct ata gga gag ttt cgt aag agg ttt ccc gga gtt ccg gtc atg gct 3216
Ala Ile Gly Glu Phe Arg Lys Arg Phe Pro Gly Val Pro Val Met Ala
1060 1065 1070

ttg aca gcg acg gca aca cag aac gtc atc ttg gat gtc aag cat aac 3264
Leu Thr Ala Thr Ala Thr Gln Asn Val Ile Leu Asp Val Lys His Asn
1075 1080 1085

ctg gca atg gag gac tgc cag act ttc tcc cag agc ttt aat cgg ccg 3312
Leu Ala Met Glu Asp Cys Gln Thr Phe Ser Gln Ser Phe Asn Arg Pro
1090 1095 1100

aac ctc tac tat gag gtc agg atg aag gag cag aat ctg att gcc cgc 3360
Asn Leu Tyr Tyr Glu Val Arg Met Lys Glu Gln Asn Leu Ile Ala Arg
1105 1110 1115 1120

atc gca gag ttg atc aag gag aag tat gac ggc cag acg ggt atc atc 3408
Ile Ala Glu Leu Ile Lys Glu Lys Tyr Asp Gly Gln Thr Gly Ile Ile
1125 1130 1135

tac aca tta tca aga aag agt gcc gag aac atc gcc aaa aat ctc cag 3456
Tyr Thr Leu Ser Arg Lys Ser Ala Glu Asn Ile Ala Lys Asn Leu Gln
1140 1145 1150

gaa aaa cac cgc atc aaa gca aag cac tac cat gcc tcc atc acc acc 3504
Glu Lys His Arg Ile Lys Ala Lys His Tyr His Ala Ser Ile Thr Thr
1155 1160 1165

gac gaa aag atc agc gtt caa cat gaa tgg caa acc ggc cga gtc aaa 3552
Asp Glu Lys Ile Ser Val Gln His Glu Trp Gln Thr Gly Arg Val Lys
1170 1175 1180

gtc gtg gta gcc acc att gcc ttc ggc atg ggc atc gac aag cct gac 3600
Val Val Val Ala Thr Ile Ala Phe Gly Met Gly Ile Asp Lys Pro Asp
1185 1190 1195 1200

gtc cgc ttt gtt atc cac cag cac atc ccc aag tcg ctc gaa ggt tac 3648
Val Arg Phe Val Ile His Gln His Ile Pro Lys Ser Leu Glu Gly Tyr
1205 1210 1215

tat caa gaa acc ggc cgc gcc gga cgt gac ggc aag cca tcg gac tgc 3696
Tyr Gln Glu Thr Gly Arg Ala Gly Arg Asp Gly Lys Pro Ser Asp Cys
1220 1225 1230

tac ttg tac ttt gcc tat ggc gac att caa tcc cta cgt cgt atg atc 3744
Tyr Leu Tyr Phe Ala Tyr Gly Asp Ile Gln Ser Leu Arg Arg Met Ile
1235 1240 1245

gcc gac ggc gaa ggt gac tac gcg caa aag gag cgt cag cta caa atg 3792
Ala Asp Gly Glu Gly Asp Tyr Ala Gln Lys Glu Arg Gln Leu Gln Met
1250 1255 1260

ctc aac cgt gtg gtc agc tat tgc gag tcg cag cac acg tgc cgg cgc 3840
Leu Asn Arg Val Val Ser Tyr Cys Glu Ser Gln His Thr Cys Arg Arg
1265 1270 1275 1280

gaa gaa gtg ctc cgc tac ttt ggc gag gag ttt gac tac cgg aag tgt 3888
Glu Glu Val Leu Arg Tyr Phe Gly Glu Glu Phe Asp Tyr Arg Lys Cys
1285 1290 1295

aga gac gga tgc gat aac tgc cgg aac gga cgc atc tcg aag tcg acg 3936
Arg Asp Gly Cys Asp Asn Cys Arg Asn Gly Arg Ile Ser Lys Ser Thr
1300 1305 1310

gag atg agg gat ttt acg gaa atc gcc ttc gcc gcg atc gag gtg gtg 3984
Glu Met Arg Asp Phe Thr Glu Ile Ala Phe Ala Ala Ile Glu Val Val
1315 1320 1325

aag agc cag cag ccc atc acg ctg ggc aag ctg tgc gac atc ctg atg 4032
Lys Ser Gln Gln Pro Ile Thr Leu Gly Lys Leu Cys Asp Ile Leu Met
1330 1335 1340

ggc aag aga aag aac gag cac ggt ggc gtg tgt cac ttt ggt atc gcc 4080
Gly Lys Arg Lys Asn Glu His Gly Gly Val Cys His Phe Gly Ile Ala
1345 1350 1355 1360

aag ggg agc acg cag agg gag ctg cag agg atc gtg ctg cag ctg aat 4128
Lys Gly Ser Thr Gln Arg Glu Leu Gln Arg Ile Val Leu Gln Leu Asn
1365 1370 1375

ttc cac aag gcg ctg ggc gag gac aat atc atg aat ggg gcg ggg atg 4176
Phe His Lys Ala Leu Gly Glu Asp Asn Ile Met Asn Gly Ala Gly Met
1380 1385 1390

cct att acc tac tat att acc ggc cct gaa gct ggt gct tac ctc tac 4224
Pro Ile Thr Tyr Tyr Ile Thr Gly Pro Glu Ala Gly Ala Tyr Leu Tyr
1395 1400 1405

aat ggc aag cgg ttg atg ctg cca gtt ccc tca aac aag tcc gtc gaa 4272
Asn Gly Lys Arg Leu Met Leu Pro Val Pro Ser Asn Lys Ser Val Glu
1410 1415 1420

ccc ccg tct cgg tct aag cag cgg agc cgt cga gtc gat gag gat atg 4320
Pro Pro Ser Arg Ser Lys Gln Arg Ser Arg Arg Val Asp Glu Asp Met
1425 1430 1435 1440

gat gag caa gaa ctt tcc acc ctg caa cga ccg cca aca tca aca aat 4368
Asp Glu Gln Glu Leu Ser Thr Leu Gln Arg Pro Pro Thr Ser Thr Asn
1445 1450 1455

gtc tct tca ccc gtt cga gcc acc aag aaa cga agt tcc aaa aag gct 4416
Val Ser Ser Pro Val Arg Ala Thr Lys Lys Arg Ser Ser Lys Lys Ala
1460 1465 1470

tta ccg acc ctc atc gcc gac tac gaa gag ccc agc tcc gac ggt cct 4464
Leu Pro Thr Leu Ile Ala Asp Tyr Glu Glu Pro Ser Ser Asp Gly Pro
1475 1480 1485

cac ggt cct ctc cac gcc aac ggc tat gag cgc gat aac ttt gtc gta 4512
His Gly Pro Leu His Ala Asn Gly Tyr Glu Arg Asp Asn Phe Val Val
1490 1495 1500

tcc gat aat gtt gaa ccc gaa gag gaa gaa gat gcc ttc gaa cct gtc 4560
Ser Asp Asn Val Glu Pro Glu Glu Glu Glu Asp Ala Phe Glu Pro Val
1505 1510 1515 1520

cgc ccc tcg cgg cgc ggc cca tct tct cgc gct acc cgc cct caa cac 4608
Arg Pro Ser Arg Arg Gly Pro Ser Ser Arg Ala Thr Arg Pro Gln His
1525 1530 1535

cgc cag acc acc ctt tat gac acc ctc tcc cac acc caa caa tcc caa 4656
Arg Gln Thr Thr Leu Tyr Asp Thr Leu Ser His Thr Gln Gln Ser Gln
1540 1545 1550

acc gtc tcc caa cac ctc gcc act ttg ggt ccg ccc atc gac gcc cgc 4704
Thr Val Ser Gln His Leu Ala Thr Leu Gly Pro Pro Ile Asp Ala Arg
1555 1560 1565

acc atg cat aac ccc cgc tac gcc cag ctt gac gag gtc cac cag gat 4752
Thr Met His Asn Pro Arg Tyr Ala Gln Leu Asp Glu Val His Gln Asp
1570 1575 1580

att gtc gat gcc ttt gtt gaa gaa gtc aag gtc ttc gag gag gac ttt 4800
Ile Val Asp Ala Phe Val Glu Glu Val Lys Val Phe Glu Glu Asp Phe
1585 1590 1595 1600

cgc aac agg aac cac atg cgc aaa ccc atc ttt acc gag acg cag tac 4848
Arg Asn Arg Asn His Met Arg Lys Pro Ile Phe Thr Glu Thr Gln Tyr
1605 1610 1615

cgt gag atg gca atc cgg tgg acg cgg tcg tta gac gcg atg cgc gcg 4896
Arg Glu Met Ala Ile Arg Trp Thr Arg Ser Leu Asp Ala Met Arg Ala
1620 1625 1630

atc ccg gat atc aac cag gat aaa gta gat cgg tat ggt gcc aaa ttc 4944
Ile Pro Asp Ile Asn Gln Asp Lys Val Asp Arg Tyr Gly Ala Lys Phe
1635 1640 1645

atc cca ctt gtg gag cgg ttc tgg ggg aat tat cag gag atg atg ggg 4992
Ile Pro Leu Val Glu Arg Phe Trp Gly Asn Tyr Gln Glu Met Met Gly
1650 1655 1660

gga ggg tat gat aat cct gct gtg gct ggc gat gag gat gat gat gag 5040
Gly Gly Tyr Asp Asn Pro Ala Val Ala Gly Asp Glu Asp Asp Asp Glu
1665 1670 1675 1680

ggc ccc agg agg aca gga aat gga aaa ggg ggg aat aag aag gga gga 5088
Gly Pro Arg Arg Thr Gly Asn Gly Lys Gly Gly Asn Lys Lys Gly Gly
1685 1690 1695

gga gga gga gga gga aat gaa gta gtg gat ttg att agt agt gat gag 5136
Gly Gly Gly Gly Gly Asn Glu Val Val Asp Leu Ile Ser Ser Asp Glu
1700 1705 1710

gat gaa cct ccg gct cgt gca cca tcg cgg aat gcg ggg cga gga aag 5184
Asp Glu Pro Pro Ala Arg Ala Pro Ser Arg Asn Ala Gly Arg Gly Lys
1715 1720 1725

gca cag tcg aca cgt ggg gga caa atc caa gat aaa ggc cga gca gtc 5232
Ala Gln Ser Thr Arg Gly Gly Gln Ile Gln Asp Lys Gly Arg Ala Val
1730 1735 1740

aac cgc cgc gga gaa ccc atc gcc gaa gaa gac gaa gaa gac tac ggg 5280
Asn Arg Arg Gly Glu Pro Ile Ala Glu Glu Asp Glu Glu Asp Tyr Gly
1745 1750 1755 1760

cta agc gac ccc gat atc gac gcc atc gat cca gac gcc atc acc gcc 5328
Leu Ser Asp Pro Asp Ile Asp Ala Ile Asp Pro Asp Ala Ile Thr Ala
1765 1770 1775

tcc gac aac tcc gac gaa gaa gat gat gat gat gat gac gaa gac ctc 5376
Ser Asp Asn Ser Asp Glu Glu Asp Asp Asp Asp Asp Asp Glu Asp Leu
1780 1785 1790

gaa tcc tcc cgc tac ttc tcc ggc tca aca ggc ccg ccc gtc tcc aaa 5424
Glu Ser Ser Arg Tyr Phe Ser Gly Ser Thr Gly Pro Pro Val Ser Lys
1795 1800 1805

gcc gtg cag gat gct cga ctc cgt gaa caa ctt tcc atg tac gcc tcc 5472
Ala Val Gln Asp Ala Arg Leu Arg Glu Gln Leu Ser Met Tyr Ala Ser
1810 1815 1820

ggc ggc agc tct tcg aaa ggt agc tac ggc tca ggg cgc gca tca gga 5520
Gly Gly Ser Ser Ser Lys Gly Ser Tyr Gly Ser Gly Arg Ala Ser Gly
1825 1830 1835 1840

gga tct tcg tcg aga gcg tcg gga tca gga tgg aga ggt gga gga gca 5568
Gly Ser Ser Ser Arg Ala Ser Gly Ser Gly Trp Arg Gly Gly Gly Ala
1845 1850 1855

ggt ggg aag aaa tac tac agg aag aag agg gct ggt tct tcg gct gct 5616
Gly Gly Lys Lys Tyr Tyr Arg Lys Lys Arg Ala Gly Ser Ser Ala Ala
1860 1865 1870

ggt ggt ggt ggt gca gga gga ggg gga gtt aca aaa cgg aag gcg agt 5664
Gly Gly Gly Gly Ala Gly Gly Gly Gly Val Thr Lys Arg Lys Ala Ser
1875 1880 1885

ggg agt ggc gcg aag acg gcg agg aag agg ggt gca tct act gcg ccg 5712
Gly Ser Gly Ala Lys Thr Ala Arg Lys Arg Gly Ala Ser Thr Ala Pro
1890 1895 1900

aag aca acg acg aga ggg gga gga tct gga gct ggg tct aga gga ggc 5760
Lys Thr Thr Thr Arg Gly Gly Gly Ser Gly Ala Gly Ser Arg Gly Gly
1905 1910 1915 1920

ggt gct ggt ggt gct ggt ggt gct ggt gct ggt gct ggt gct gga gga 5808
Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Ala Gly Ala Gly Gly
1925 1930 1935

ggg aaa agg ggt ggt gga ggt gga gga gga atg gga ggg ata agt gtt 5856
Gly Lys Arg Gly Gly Gly Gly Gly Gly Gly Met Gly Gly Ile Ser Val
1940 1945 1950

atg cct cat tag 5868
Met Pro His
1955

<210> 3
<211> 1955
<212> PRT
<213> Neurospora crassa

<400> 3
Met Ala Lys Leu Ser Val Lys Asn Asn Leu Pro Arg Pro His Leu Val
1 5 10 15
Ser Leu Ser Ser Ser Thr Thr Gly Ser Gly Ser Gly Ser Ala Ser Arg
20 25 30
Ser Ala Ser Ala Lys His Gly Ser Ala Gly Ser Ser Thr Phe Asp His
35 40 45
Glu Gln His Gln Gln His Gln Gln Gln Gln Gln Gln Lys Arg Gln Arg
50 55 60
Ser Gln Ser Glu Ala Arg Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
65 70 75 80
Gln Gln Gln Gln Gln Gln Gln Gln Ala Gln His His Ala His Ser Thr
85 90 95
Tyr Ala Gln Arg Pro Gln Pro Thr Pro Gln Gln Arg Pro Pro Gln Asn
100 105 110
Leu Leu Thr Pro Ala Ser Thr Thr Gly Ala Ser Val Gly Pro Leu Gln
115 120 125
Arg Ala Tyr Ser Val Ser Leu Ala Ala Arg Gln Ser Pro Ser Thr Asn
130 135 140
Leu Val Arg Pro Lys Thr Asp Ser Pro Ala Pro His Thr Leu His Leu
145 150 155 160
Lys Asn Lys Lys Asn Leu Arg His Pro Ala Pro Thr Pro Asp Ser Pro

165 170 175
Ile Val Asp Asp Asp Ile Phe Ser Asp Ala Val Asp Leu Thr Glu Glu
180 185 190
Leu Asp His Asp His Asp Leu Asn Gly Lys Asp Lys Asp Asn Thr Asp
195 200 205
Asn Asp Asn Thr Val Ala Ser Ser Ser Leu Ile Gly Phe Gly Asp Asp
210 215 220
Lys Leu Leu Trp Arg Glu Asp Phe Ala Glu Arg Ala Glu Pro Glu His
225 230 235 240
Glu Arg Gly Gly Ser Arg Pro Arg Gln Val Lys Lys Arg Lys Ile Ser
245 250 255
Asn Asp Tyr Ile Met Lys Asp Glu Asp Val Ser Leu Phe Asp Asp Asp
260 265 270
Gly Glu Glu Asp Glu Phe Met Asp Ile Asn Glu Leu Val Gln Gly Asp
275 280 285
Arg Glu Ser Thr Pro Lys Pro Lys Ala Thr Ser Arg Ser Val Ser Thr
290 295 300
Arg Leu Pro Pro Thr Val Ser Leu Gln Arg Gly Arg Ser Pro Lys Arg
305 310 315 320
Lys Glu Ala Ser Val Glu Lys Arg Thr Thr Glu Asn Gln Gln Gln Ala
325 330 335
Asp Arg Glu Asp Glu Pro Ser Phe Met Ser Ser Pro Asp Val Asp Asn
340 345 350
Ser Arg Lys Arg Lys Ser Ser Gly Ser Pro Thr Gly Leu Thr Thr Pro
355 360 365
Arg Pro Gln Gln Lys Gln Thr Glu Glu Val Pro Gly Thr Thr Thr Ala
370 375 380
Lys Lys Pro Arg Arg Ser Glu Val Met Asp Ser Glu Asp Glu Ala Phe
385 390 395 400
Thr Pro Leu Ser Ala Gly Ser Leu Pro Gly Ser Ala Glu Phe Phe Arg
405 410 415
Ser Gly Gly Thr Thr Thr Arg Glu Leu Gly Leu Asp Glu Asp Thr Val
420 425 430
Met Asp Thr Pro Ser Arg Pro Pro Val Glu Ser Thr Leu Pro Thr Leu
435 440 445
Glu Ser Val Glu Ser Arg Pro Pro Pro Leu Pro Pro Met Asp Leu Pro
450 455 460
Ser Gln Arg Lys Pro Leu Glu Pro Leu Asn Thr Pro Arg Asn Gln Leu
465 470 475 480
Leu Glu Ser Val Glu Arg Pro Thr Gln Gln Pro Ser Val Gly Pro Ser
485 490 495
Phe Ala Gln Ser Ser Thr Leu Ala Glu Ser Ser Leu Pro Pro Ser Met
500 505 510
Pro Pro Pro Ser Glu Asp Pro Leu Asn Thr Arg Glu Asn Ser Asn Leu
515 520 525
Glu Glu Phe Asp Tyr Lys Leu Tyr Lys Pro Leu Leu Asp Leu Phe Val
530 535 540
Asn Ala Pro Ala Ile Leu Glu Arg Glu Leu Ser Ala Val Asn Asp Glu

545 550 555 560
Leu Gln Glu Asn Met Ile Lys Leu Arg Asp Cys Leu Arg Leu Pro Arg
565 570 575
Glu Glu Arg Asp Arg Ala Arg Glu Glu Val Lys Lys Glu Lys Glu Met
580 585 590
Leu Lys Arg Arg Asp Ile Ala Leu Arg Ala Leu Gln Asp Glu His Lys
595 600 605
Leu Tyr Val Lys Lys Arg Lys Glu His Asn Leu Ile Asn Glu Glu Ile
610 615 620
Val Arg Ala Tyr Ala Glu Glu Asp Asp Glu Tyr Glu Asp Gln Leu Met
625 630 635 640
Ala Gln Leu Asp Lys Leu Asp Asp Glu Val Glu Ala Ile Val Lys Ser
645 650 655
Leu Thr Arg Leu Ile Val Ala Ala Gly Ile Thr Glu Lys Ser Phe Asp
660 665 670
Leu Lys Lys Glu Glu Glu Glu Glu Glu Glu Lys Pro Ile Ile Ile Ala
675 680 685
Thr Pro Thr Pro Ser Thr Arg Thr Glu Ala Pro Val Leu Pro Thr Thr
690 695 700
Glu Tyr His Asn Ser Gln Gln Val Ile Leu Gln Thr Gln His Pro Ala
705 710 715 720
Ala Gln Gln Val Ser His Arg Val Pro Pro Pro Pro Thr Pro Ser Phe
725 730 735
Gln Thr Ala Arg Gln Thr Pro Val Ser Tyr Gln Ser Arg Pro Thr Asn
740 745 750
Asn Ser Phe Pro Asp Ile Ser Ala Glu Glu Ala Met Met Phe Asp Lys
755 760 765
Glu Asp Pro Phe Met Glu Gln Gln His Ala Pro Ala Ser Ala Pro Phe
770 775 780
Gln Ala Thr Leu Pro Gln Arg Asn Ser Pro Phe Lys Thr Ala Pro Phe
785 790 795 800
Lys Pro Val His Gly His Asp Tyr Phe Asp Asp Glu Asp Asp Asp Ala
805 810 815
Asp Leu Leu Ala Ala Val Asp Ser Ala Glu Thr Tyr Thr Ser Thr Ala
820 825 830
Ala Thr Thr Thr Thr Asn Asn Asn Asn His Leu Arg Ser Gln Ser Val
835 840 845
Met Ser Thr Ser Thr Ala Thr Thr Ile Lys Pro Arg Lys Arg Asn Glu
850 855 860
Asn Ala Asn Ala Lys Lys Pro Lys Ser Val His Ala Lys Leu Ser Met
865 870 875 880
Pro Pro Glu Lys Met Lys Tyr Ala Trp Ser Asn Asp Val Arg Lys Ala
885 890 895
Leu Lys Asp Arg Phe Arg Met Ser Gly Phe Arg Gln Asn Gln Leu Glu
900 905 910
Ala Ile Asn Ala Thr Leu Gly Gly Lys Asp Ala Phe Val Leu Met Pro
915 920 925
Thr Gly Gly Gly Lys Ser Leu Cys Tyr Gln Leu Pro Ala Val Val Arg

930 935 940
Ser Gly Lys Thr Arg Gly Ile Thr Val Val Ile Ser Pro Leu Leu Ser
945 950 955 960
Leu Met Leu Asp Gln Val Asn His Leu Ala Asn Leu Met Ile Gln Ala
965 970 975
Tyr Ala Phe Asn Gly Asp Met Asn Ser Glu Met Arg Arg Met Val Phe
980 985 990
Gln Lys Leu Asp Ala Glu His Pro Glu His Glu Leu Gln Leu Leu Tyr
995 1000 1005
Val Thr Pro Glu Met Val Ser Lys Asn Gln Thr Phe Val Asn Lys Met
1010 1015 1020
Met Asp Leu Tyr Arg Arg Lys Lys Leu Ala Arg Ile Val Ile Asp Glu
1025 1030 1035 1040
Ala His Cys Val Ser Gln Trp Gly His Asp Phe Arg Pro Asp Tyr Lys
1045 1050 1055
Ala Ile Gly Glu Phe Arg Lys Arg Phe Pro Gly Val Pro Val Met Ala
1060 1065 1070
Leu Thr Ala Thr Ala Thr Gln Asn Val Ile Leu Asp Val Lys His Asn
1075 1080 1085
Leu Ala Met Glu Asp Cys Gln Thr Phe Ser Gln Ser Phe Asn Arg Pro
1090 1095 1100
Asn Leu Tyr Tyr Glu Val Arg Met Lys Glu Gln Asn Leu Ile Ala Arg
1105 1110 1115 1120
Ile Ala Glu Leu Ile Lys Glu Lys Tyr Asp Gly Gln Thr Gly Ile Ile
1125 1130 1135
Tyr Thr Leu Ser Arg Lys Ser Ala Glu Asn Ile Ala Lys Asn Leu Gln
1140 1145 1150
Glu Lys His Arg Ile Lys Ala Lys His Tyr His Ala Ser Ile Thr Thr
1155 1160 1165
Asp Glu Lys Ile Ser Val Gln His Glu Trp Gln Thr Gly Arg Val Lys
1170 1175 1180
Val Val Val Ala Thr Ile Ala Phe Gly Met Gly Ile Asp Lys Pro Asp
1185 1190 1195 1200
Val Arg Phe Val Ile His Gln His Ile Pro Lys Ser Leu Glu Gly Tyr
1205 1210 1215
Tyr Gln Glu Thr Gly Arg Ala Gly Arg Asp Gly Lys Pro Ser Asp Cys
1220 1225 1230
Tyr Leu Tyr Phe Ala Tyr Gly Asp Ile Gln Ser Leu Arg Arg Met Ile
1235 1240 1245
Ala Asp Gly Glu Gly Asp Tyr Ala Gln Lys Glu Arg Gln Leu Gln Met
1250 1255 1260
Leu Asn Arg Val Val Ser Tyr Cys Glu Ser Gln His Thr Cys Arg Arg
1265 1270 1275 1280
Glu Glu Val Leu Arg Tyr Phe Gly Glu Glu Phe Asp Tyr Arg Lys Cys
1285 1290 1295
Arg Asp Gly Cys Asp Asn Cys Arg Asn Gly Arg Ile Ser Lys Ser Thr
1300 1305 1310
Glu Met Arg Asp Phe Thr Glu Ile Ala Phe Ala Ala Ile Glu Val Val

1315 1320 1325
Lys Ser Gln Gln Pro Ile Thr Leu Gly Lys Leu Cys Asp Ile Leu Met
1330 1335 1340
Gly Lys Arg Lys Asn Glu His Gly Gly Val Cys His Phe Gly Ile Ala
1345 1350 1355 1360
Lys Gly Ser Thr Gln Arg Glu Leu Gln Arg Ile Val Leu Gln Leu Asn
1365 1370 1375
Phe His Lys Ala Leu Gly Glu Asp Asn Ile Met Asn Gly Ala Gly Met
1380 1385 1390
Pro Ile Thr Tyr Tyr Ile Thr Gly Pro Glu Ala Gly Ala Tyr Leu Tyr
1395 1400 1405
Asn Gly Lys Arg Leu Met Leu Pro Val Pro Ser Asn Lys Ser Val Glu
1410 1415 1420
Pro Pro Ser Arg Ser Lys Gln Arg Ser Arg Arg Val Asp Glu Asp Met
1425 1430 1435 1440
Asp Glu Gln Glu Leu Ser Thr Leu Gln Arg Pro Pro Thr Ser Thr Asn
1445 1450 1455
Val Ser Ser Pro Val Arg Ala Thr Lys Lys Arg Ser Ser Lys Lys Ala
1460 1465 1470
Leu Pro Thr Leu Ile Ala Asp Tyr Glu Glu Pro Ser Ser Asp Gly Pro
1475 1480 1485
His Gly Pro Leu His Ala Asn Gly Tyr Glu Arg Asp Asn Phe Val Val
1490 1495 1500
Ser Asp Asn Val Glu Pro Glu Glu Glu Glu Asp Ala Phe Glu Pro Val
1505 1510 1515 1520
Arg Pro Ser Arg Arg Gly Pro Ser Ser Arg Ala Thr Arg Pro Gln His
1525 1530 1535
Arg Gln Thr Thr Leu Tyr Asp Thr Leu Ser His Thr Gln Gln Ser Gln
1540 1545 1550
Thr Val Ser Gln His Leu Ala Thr Leu Gly Pro Pro Ile Asp Ala Arg
1555 1560 1565
Thr Met His Asn Pro Arg Tyr Ala Gln Leu Asp Glu Val His Gln Asp
1570 1575 1580
Ile Val Asp Ala Phe Val Glu Glu Val Lys Val Phe Glu Glu Asp Phe
1585 1590 1595 1600
Arg Asn Arg Asn His Met Arg Lys Pro Ile Phe Thr Glu Thr Gln Tyr
1605 1610 1615
Arg Glu Met Ala Ile Arg Trp Thr Arg Ser Leu Asp Ala Met Arg Ala
1620 1625 1630
Ile Pro Asp Ile Asn Gln Asp Lys Val Asp Arg Tyr Gly Ala Lys Phe
1635 1640 1645
Ile Pro Leu Val Glu Arg Phe Trp Gly Asn Tyr Gln Glu Met Met Gly
1650 1655 1660
Gly Gly Tyr Asp Asn Pro Ala Val Ala Gly Asp Glu Asp Asp Asp Glu
1665 1670 1675 1680
Gly Pro Arg Arg Thr Gly Asn Gly Lys Gly Gly Asn Lys Lys Gly Gly
1685 1690 1695
Gly Gly Gly Gly Gly Asn Glu Val Val Asp Leu Ile Ser Ser Asp Glu

1700 1705 1710

Asp Glu Pro Pro Ala Arg Ala Pro Ser Arg Asn Ala Gly Arg Gly Lys
1715 1720 1725

Ala Gln Ser Thr Arg Gly Gly Gln Ile Gln Asp Lys Gly Arg Ala Val
1730 1735 1740

Asn Arg Arg Gly Glu Pro Ile Ala Glu Glu Asp Glu Glu Asp Tyr Gly
1745 1750 1755 1760

Leu Ser Asp Pro Asp Ile Asp Ala Ile Asp Pro Asp Ala Ile Thr Ala
1765 1770 1775

Ser Asp Asn Ser Asp Glu Glu Asp Asp Asp Asp Asp Asp Glu Asp Leu
1780 1785 1790

Glu Ser Ser Arg Tyr Phe Ser Gly Ser Thr Gly Pro Pro Val Ser Lys
1795 1800 1805

Ala Val Gln Asp Ala Arg Leu Arg Glu Gln Leu Ser Met Tyr Ala Ser
1810 1815 1820

Gly Gly Ser Ser Ser Lys Gly Ser Tyr Gly Ser Gly Arg Ala Ser Gly
1825 1830 1835 1840

Gly Ser Ser Ser Arg Ala Ser Gly Ser Gly Trp Arg Gly Gly Gly Ala
1845 1850 1855

Gly Gly Lys Lys Tyr Tyr Arg Lys Lys Arg Ala Gly Ser Ser Ala Ala
1860 1865 1870

Gly Gly Gly Gly Ala Gly Gly Gly Gly Val Thr Lys Arg Lys Ala Ser
1875 1880 1885

Gly Ser Gly Ala Lys Thr Ala Arg Lys Arg Gly Ala Ser Thr Ala Pro
1890 1895 1900

Lys Thr Thr Thr Arg Gly Gly Gly Ser Gly Ala Gly Ser Arg Gly Gly
1905 1910 1915 1920

Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Ala Gly Ala Gly Gly
1925 1930 1935

Gly Lys Arg Gly Gly Gly Gly Gly Gly Gly Met Gly Gly Ile Ser Val
1940 1945 1950

Met Pro His
1955

International Application No
PCT/IT 99/00391

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/31 C12N15/63 C12N15/67 C12N15/70 C12N15/74 C12N15/80 C12N15/82 C12N15/85 C12N15/11 C12N9/90 C12N1/19 C12N1/21 C12N5/10 C07K14/37 A01K67/027

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 17979 A (NEW YORK BLOOD CENTER INC) 22 May 1997 (1997-05-22) 47.2% identity in 439 aa overlap with amino acids 897-1330 of SeqIdNo.3	1,2,7, 10,11, 16,18,19
X	-& DATABASE GENESEQ EBI, Hinxton, U.K. Accession Number: W31551, 27 January 1998 (1998-01-27) ELLIS N ET AL: "Bloom's syndrome BLM mutated protein" XP002136373 47.2% identity in 439 aa overlap with amino acids 897-1330 of SeqIdNo.3 abstract --- -/--	1,2,18, 19

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 April 2000	11/05/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Lonnoy, O

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	COGONI C ET AL: "Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., vol. 94, no. 19, 16 September 1997 (1997-09-16), pages 10233-10238, XP002136370 cited in the application table 2	1-23
Y	LINDEN H ET AL: "White collar 2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in Neurospora crassa" EMBO JOURNAL., vol. 16, no. 1, 2 January 1997 (1997-01-02), pages 98-109, XP002136371 page 99, column 1, paragraph 3 -page 100, column 1, paragraph 2	1-23
A	SHERMAN J M ET AL: "An uncertain silence" TRENDS IN GENETICS,NL,ELSEVIER SCIENCE PUBLISHERS B.V. AMSTERDAM, vol. 13, no. 8, 1 August 1997 (1997-08-01), pages 308-313, XP004084604 ISSN: 0168-9525	
A	COGONI C ET AL: "Quelling: transgene-induced gene silencing in Neurospora crassa" NATO ADVANCED STUDY INSTITUTE SERIES, SERIES H CELL BIOLOGY, vol. 104, 1998, pages 103-112, XP000906708	
T	COGONI C AND MACINO G: "Posttranscriptional gene silencing in Neurospora by a RecQ DNA helicase" SCIENCE., vol. 286, no. 5448, 17 December 1999 (1999-12-17), pages 2342-2344, XP002136372 the whole document	1-24

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9717979 A	22-05-1997	US 5824501 A	20-10-1998
		CA 2237356 A	22-05-1997
		EP 0862442 A	09-09-1998